

Université Montpellier II  
Ecole Doctorale Systèmes Intégrés  
en Biologie, Agronomie, Géosciences, Hydrosiences, Environnement

**NOTICE DE TRAVAUX**  
**Présentée pour l'obtention du diplôme de :**

**HABILITATION A DIRIGER DES RECHERCHES**

Par  
Thierry Leroy  
Chercheur au Cirad

**Amélioration des caféiers et de la qualité des cafés,  
Du champ aux gènes**

Soutenue le 4 juillet 2008

Membres du jury :

Didier Courtois	DR	Centre de Recherches Nestlé Tours	Examineur
Françoise Dosba	PR	SupAgro, Montpellier	Examineur
Pascal Gantet	PR	Université Montpellier II	Examineur
Maria Manzanares-Dauleux	MC	ENSAR/Agrocampus Rennes	Examineur
Jean-Claude Pech	PR	INRA/INP-ENSAT	Rapporteur
Gilles Pilate	DR	INRA	Rapporteur
Christophe Plomion	DR	INRA	Rapporteur

## TABLE DES MATIERES

I. CURRICULUM VITAE	3
II ACTIVITES D'ENCADREMENT	5
III PUBLICATIONS	8
IV. SYNTHESE DES RESULTATS SCIENTIFIQUES	13
Introduction	13
A Analyse de la diversité	14
1 Evaluation de la diversité de <i>C. canephora</i> en Côte d'Ivoire	14
2 Evaluation de la diversité de <i>C. canephora</i> en Ouganda	22
3 Diversité de <i>C. arabica</i> et autres espèces	26
B Transmission des caractères et gains génétiques par sélection classique	28
C Amélioration des caféiers par transfert de gènes	32
1 Quelles constructions génétiques ?	33
2 Obtention des plantes transformées	34
3 Evaluation des plantes	34
4 L'essai au champ	37
5 Partenariats et perspectives	37
D Génomique de la qualité des cafés	39
1 Mise en place des outils moléculaires	40
2 Mécanismes moléculaires de la construction de la qualité	41
3 Carte génétique de <i>C. canephora</i>	44
4 Diversité moléculaire	47
5 Analyses préliminaires du DL dans des populations de <i>C. canephora</i>	51
Conclusions	52
V LE PROJET DE RECHERCHES	54
A. Les questions de recherche	54
B. Contexte - L'amélioration des caféiers	54
C. Analyses de QTL	56
D. Analyses de DL	56
E. Etudes d'association	58
1 Quelles populations de travail?	58
2 Evaluation génotypique des collections	60
3 Analyses phénotypiques	61
4 Analyse des associations	61
F. Perspectives de sujets de recherche	62
Références bibliographiques citées dans le document	64

# I. CURRICULUM VITAE

**Thierry Leroy**

Né le 01 novembre 1960

**Généticien café au Cirad**

Cirad-BIOS, TA A 96/03, 34398 Montpellier Cedex 5

## **Formation académique :**

Thèse de l'ENSA de Rennes, 1993. Mention Biologie et Agronomie. Titre : Diversité, paramètres génétiques et amélioration par sélection récurrente réciproque du caféier *Coffea canephora* P.

Ingénieur Agronome ENSA de Rennes, 1983.

## **Formation complémentaire :**

Suivi DEA « Bases de la production végétale », Université Montpellier, Pr Macheix, 1993.

## **Champ de recherches**

1. Analyse du génome
2. Diversité et amélioration génétique
3. Création, analyse, mise au champ et évaluation de plantes transgéniques.

## **Expériences-clés**

- 10 ans de recherches en amélioration des plantes au champ sur le caféier *C. canephora*: gestion et analyse de la diversité, génétique quantitative,
- 7 ans de recherches sur la transgénèse chez les caféiers: culture in vitro, transformation génétique, analyse moléculaire, entomologie, analyse protéique, comportement agronomique,
- 7 ans de recherches en analyse du génome: diversité moléculaire, génomique structurale et fonctionnelle, marquage moléculaire,
- 16 ans de gestion de la recherche: animation et gestion d'une équipe, coordination de réseaux, conception et suivi de projets, accueil d'étudiants et de chercheurs (1989-1992 ; 1994-2007).

## **Expérience professionnelle :**

**2007- :** Cirad-BIOS, UMR DAP, équipe Structure des Ressources génétiques, en charge des recherches menées sur le caféier (construction de la qualité, cartographie génétique et physique du caféier, recherche de gènes), coordinateur des recherches menées sur le caféier dans l'ensemble des équipes de l'UMR.

**2005-2006** : Cirad, département des cultures pérennes, UMR PIA, équipe café, responsable de l'équipe café de l'UMR impliquée dans des travaux sur la génomique et la génétique des caféiers, pour l'amélioration de la qualité et la tolérance à la sécheresse.

**2003-2004** : Cirad, département Cultures Pérennes, programme café, responsable du thème sur l'amélioration des caféiers pour leur production, en quantité et en qualité. Les travaux sont essentiellement orientés autour de la génétique classique et de la cartographie génétique, avec la création de ressources génomiques comme les marqueurs microsatellites et la banque génomique de clones BAC. Les travaux de physiologie moléculaire sont initiés en 2004 en collaboration avec des équipes brésiliennes.

**2000-2002** : Cirad, département CP, programme café, localisé dans les laboratoires de l'IRD, responsable du thème 4 du programme café (contrôle génétique et piégeage du scolyte des baies du caféier), impliqué dans des travaux de génétique, de transformation génétique et de recherche de gènes, mais aussi de contrôle par piégeage des populations de scolyte des fruits. Création, avec l'IRD, d'une équipe commune sur la génétique et la génomique des caféiers (l'équipe commune prend fin en septembre 2002).

**1994-2000** : Cirad, département Cultures pérennes, détaché sur le centre de recherches Nestlé de Tours dans le cadre d'un programme conjoint sur la transformation génétique des caféiers, responsabilité de ce programme conjoint Cirad/Nestlé sur la modification génétique des caféiers. Développement des méthodologies de transformation génétique des caféiers, introduction dans les caféiers de gènes de résistance à la mineuse des feuilles, mise en place et évaluation d'un essai de caféiers génétiquement modifiés en Guyane française. A partir de 1998, animation du thème 4 du programme café.

**1992-1994** : Cirad, département des cultures pérennes. Rédaction et présentation de la thèse de Doctorat. Formation en culture in vitro, biologie moléculaire et transformation génétique en vue de la mise en place du programme commun avec Nestlé.

**1989-1992** : Cirad, Institut de recherches sur le café et cacao (IRCC), localisé à l'IDFOR-DCC (Côte d'Ivoire, Bingerville). Généticien, en charge de l'ensemble des programmes d'amélioration génétique des caféiers en Côte d'Ivoire, en particulier des essais localisés dans les trois stations principales de recherche à Bingerville, Divo et Abengourou.

**1984-1989** : CIRAD-IRCC, Divo, Côte d'Ivoire. Généticien, responsable de la mise en place d'un programme de sélection réciproque sur le caféier en Côte d'Ivoire : évaluation des collections, introduction et évaluation de nouveaux génotypes, réalisation de croisements, sélection des arbres et amélioration des populations en sélection.

## **Publications**

30 publications dans des revues à comité de lecture,  
6 chapitres d'ouvrages,  
53 communications à des réunions internationales.

## **Responsabilités internationales**

Membre fondateur, responsable du Working Group « marqueurs et cartographie génétique » et membre du steering committee de l'ICGN, International Coffee Genome Network.

## II ACTIVITES D'ENCADREMENT

### *Thèses*

**Philippe Cubry (2005-2008)**, Université Montpellier II, Ecole doctorale SIBAGHE. Thèse en cours, soutenance prévue en septembre 2008.

Sujet : Structuration de la diversité génétique et analyse des patrons de déséquilibre de liaison de populations de l'espèce *Coffea canephora*.

Ces travaux permettent d'analyser l'ensemble de la variabilité de l'espèce *C. canephora* et d'analyser les déséquilibres de liaison pour quelques populations dans des régions d'intérêt spécifiques du génome, pour lesquelles la densité de marqueurs cartographiés est forte.

Codirecteur de la thèse avec Claire Lanaud.

Publications : 1 publication dans Genome, une autre en cours pour Tree, Genetics & Genome.

**Pascal Musoli (2004-2007)**, SupAgro Montpellier, Ecole doctorale SIBAGHE. Thèse présentée en mars 2007.

Titre : Recherche de sources de résistance à la trachéomycose du caféier *Coffea canephora* Pierre, due à *Fusarium xylarioides* Steyaert en Ouganda.

Ces travaux ont permis de mieux connaître les mécanismes de diffusion et de transmission de cette maladie létale des caféiers, et de mettre en évidence la grande variabilité génétique des caféiers d'Ouganda, spontanés ou cultivés, y compris pour la tolérance à cette maladie.

Participation à l'encadrement avec Daniel Bieysse et Christian Cilas (Chercheurs Cirad).

Participation au comité de thèse, membre du jury de la thèse.

Pascal Musoli est maintenant responsable de la recherche sur l'amélioration des caféiers au NARO, Institut national de recherche en Ouganda.

Publications : 1 publication en cours de soumission à Genome, 1 autre en préparation pour Euphytica.

### *Master 2 ou DEA*

**Philippe Cubry (2005)**, DEA « Ressources phytogénétiques et interactions biologiques », UM2, 2005.

Titre : Analyse de la diversité et évaluation du déséquilibre de liaison chez quelques populations naturelles et cultivées de caféiers *Coffea canephora*.

Co-encadrement avec Magali Dufour (Cirad).

Philippe Cubry effectue sa thèse de Doctorat au sein de l'équipe, toujours sur les caféiers.

**Sophie Bouchet (2005)**, Master 2 recherche STS, mention biologie, spécialité : génétique, adaptation et productions végétales, Université/Agrocampus Rennes, 2005.

Titre : diversité nucléotidique et évolution moléculaire de cinq gènes impliqués dans la voie de biosynthèse du saccharose au sein du genre *Coffea*.

Co-encadrement avec David Pot (Cirad).

Sophie Bouchet effectue sa thèse de Doctorat au Cirad sur des populations de sorgho.

Titre : Génomique des populations chez les plantes cultivées : exemple du sorgho. Sa thèse est effectuée sous la Direction de Jean-Christophe Glaszmann.

**Komlan Avia (2004)**, DEA de Génétique, adaptation et productions végétales.

Université/Agrocampus Rennes, 2004

Titre : Cartographie génétique de *Coffea canephora* et premières analyses du déséquilibre de liaison dans onze populations.

Encadrement.

Komlan Avia effectue sa thèse de Doctorat à l'INRA de Estrées-Mons (80) sur la résistance au gel des légumineuses. Titre : cartographie de QTL biochimiques et physiologiques de l'acclimatation au froid chez *Medicago truncatula* et approche gènes candidats. Sa thèse est effectuée sous la Direction de Isabelle Lejeune Hénaut.

**Christophe Montagnon (1989)**, DAA, Stage de fin d'études de l'ENSA de Rennes pour l'obtention du DAA, 1989.

Titre : contribution à l'évaluation de la collection de *Coffea canephora* de l'IRCC en relation avec le schéma de sélection récurrente et réciproque.

Encadrement.

Christophe Montagnon a été recruté au Cirad en 1989, et a travaillé dans l'équipe sur l'amélioration des caféiers en Côte d'Ivoire jusqu'en 1992. Il a ensuite travaillé dans d'autres pays, avant de revenir en Côte d'Ivoire à la fin des années 90. Il est actuellement en poste au Mexique dans le cadre d'un partenariat entre le Cirad et une entreprise, pour l'amélioration des caféiers en Amérique centrale et du Sud.

Publications : 19 publications communes entre 1991 et 2008.

### ***DESS ou Master 1***

**Noël Durand (2006)**, Master 1 BGAE spécialité Biodiversité, biodétection Bio traçabilité, 2006.

Titre : variabilité biochimique et nucléotidique du métabolisme des diterpènes dans les génotypes de *Coffea canephora* d'Ouganda.

Co encadrement avec Bernard Guyot (Cirad).

Noël Durand a effectué son Master2 3B à l'Université de Montpellier, rôle d'un éliciteur sur le développement du CBD. Il doit démarrer une thèse fin 2008.

**Margot Berline (2002)**, DESS Bioingénierie, option : biotechnologies végétales. Université Toulouse 3, 2002.

Titre : Définition de marqueurs microsatellites polymorphes et cartographie génétique préliminaire d'une descendance intraspécifique de *Coffea canephora*.

Co-encadrement avec Magali Dufour (Cirad).

Margot Berline a réorienté ses activités vers le management de la recherche après son DESS. Après avoir suivi un Mastère en management de l'innovation scientifique, elle travaille maintenant comme coordinatrice du pôle Atlantic Biotherapies chez Atlanpôle.

### ***Post-doctorants***

**Monique Royer (1994-1995)**, Cirad Montpellier.

Titre des travaux : Construction de vecteurs binaires pour la création de caféiers transgéniques résistants à *Perileucoptera coffeella*.

Co-encadrement avec R. Frutos.

1 Publication dans Plant Cell Reports.

Monique Royer travaille au Cirad dans l'UMR BGPI sur les stress biotiques chez la canne à sucre, elle est responsable d'une équipe de recherches.

**Anne-Marie Henry (1998)**, Nestlé Tours.

Titre des travaux : analyse moléculaire des plantes génétiquement modifiées produites par le laboratoire.

Encadrement.

1 Publication dans Plant Cell Reports

Anne-Marie Henry est Maître Assistant à l'Université de Perpignan.

### ***Autres encadrements***

**Pauline Aluka**, Cadre de recherches ougandais, deux stages de 6 mois en **2004 et 2005**.

Titre des travaux : projet USDA/ICRAF, Robusta quality markers in Africa.

Co-encadrement avec Fabrice Davrieux et Magali Dufour (Cirad).

Publications : rapports du projet, publication en cours de rédaction.

Pauline Aluka est chercheur sur le café au CORI (Coffee Research Institute) en Ouganda.

**Louis-Marie Raboin (1994)**. Université Montpellier, Certificat de spécialisation informatique.

Titre : Conception d'une base de données sous foxpro2.5 DOS.

Encadrement.

Formatage et mise en place de la base de données des collections de caféiers de Côte d'Ivoire avec le logiciel Foxpro2.5.

Louis-Marie Raboin est chercheur au Cirad et travaille à Madagascar.

**Franck Bodin (1988)**.ISTOM.

Stage ISTOM deuxième année de trois mois sur la station de Divo.

Encadrement.

A participé à la mise en place de la base de données des clones.

**F. Souza Pereira, M. Dias, M. Luzolo, R. Mayimona, mai-juin 1986**. Stagiaires angolais sur la station de Divo.

Titre du travail : sélection des caféiers écimés en collection à Divo.

Encadrement.

**Christophe Duchemin (1986)**.ISTOM.

stage ISTOM de deuxième année de trois mois sur la station de Divo.

Encadrement.

A participé à l'évaluation phénotypique de la collection de caféiers de la station de Divo.

### ***Activités d'enseignement***

Cours spécialisés sur les plantes tropicales (2<sup>ème</sup> année, ENSA de Rennes) : 1996-1998, 4 heures.

Module DEA Université de Montpellier sur les contraintes de la mise au champ des plantes transgéniques : 2001-2005 (2,5 heures).

### III PUBLICATIONS

29 articles publiés ou sous presse dans des revues à comité de lecture, 1 article soumis, 6 chapitres d'ouvrage, deux mémoires (53 communications ou posters dans des congrès internationaux non comptabilisés).

<b>Journal</b>	<b>Nombre</b>	<b>Dont en Premier auteur</b>	<b>Dont en Dernier auteur</b>	<b>Indice d'impact (2006)</b>
Agronomie	1			0.86
Annals of Applied biology	1			1.38
Euphytica	7	3	1	0.91
Genome	3		3	1.97
International Journal of Pest management	1			0.71
Journal of Experimental Botany	1			3.63
Molecular Genetics and genomics	1			2.55
Plant Breeding	1		1	0.95
Plant Cell Reports	1	1		1.73
Plant Physiology and Biochemistry	1			1.85
Theoretical and Applied Genetics	1	1		2.72
Tropical Science	1			---
Brazilian Journal of Plant Physiology	1	1		---
Café Cacao Thé	6	2	1	Revue disparue
Plantations, Recherche, Développement	3			Revue disparue



### Articles dans des revues à comité de lecture

Geromel C, Ferreira LP, Bonatelli, Bottcher A, Pot D, Pereira LFP, **Leroy T**, Vieira LGE, Mazzafera P, Marraccini P (2008). Sucrose metabolism in *Coffea racemosa* endosperm during fruit development. *Annals of Applied Biology*, 152: 179-187.

Geromel C, Ferreira LP, Davrieux F, Guyot B, Ribeyre F, dos Santos Scholz MB, Pereira LFP, Vaast P, Pot D, **Leroy T**, Androcioli Filho A, Vieira LGE, Mazzafera P, Marraccini P (2008). Effects of shade on the development and sugar metabolism of coffee (*Coffea arabica*) fruits. *Plant physiology and Biochemistry*, 46: 569-579.

*Musoli P, Cubry P, Aluka P, Billot C, Dufour M, De Bellis F, Pot D, Bieysse D, Charrier A, **Leroy T**. Genetic differentiation of wild and cultivated populations: diversity of Coffea canephora in Uganda. Soumis à Genome.*

Cubry P, Musoli P, Legnaté H, Pot D, De Bellis F, Poncet V, Anthony F, Dufour M, **Leroy T** (2008). Diversity in coffee using SSR markers: structure of the *Coffea* genus and perspectives for breeding. *Genome*, 51 (1): 50-63.

Montagnon C., **Leroy T.**, Cilas C., Legnaté H., Charrier A (2008). Heterozygous genotypes are efficient testers for assessing between-population combining ability in the reciprocal recurrent selection of *Coffea canephora*. *Euphytica*, 160: 101-110.

Bustamante\_Porras J, Campa C, Poncet V, Noirot M, **Leroy T**, Hamon S, de Kochko A (2007). Molecular Characterization of an Ethylene Receptor gene (*CcETR1*) in coffee trees, its relationships with fruit development and caffeine content. *Molecular Genetics and Genomics*, 277: 701-712.

Poncet V, Dufour M, Hamon P, Hamon S, De Kochko A, **Leroy T** (2007). Development of genomic microsatellite markers in *Coffea canephora* and their transferability to other coffee species. *Genome* 50(12): 1156-1161.

Geromel C, Ferreira LP, Guerreiro SMC, Cavalari AA, Pot D, Pereira LFP, **Leroy T**, Vieira LGE, Mazzafera P, Marraccini P (2006). Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. *Journal of Experimental Botany*, 57 (12): 3243-3258.

**Leroy T**, Ribeyre F, Bertrand B, Charmetant P, Dufour M, Montagnon C, Marraccini P, Pot D (2006). Genetics of coffee quality. *Brazilian Journal of Plant Physiology*, 18(1): 229-242.

Perthuis B, Pradon JL, Montagnon C, Dufour M, **Leroy T** (2005). Stable resistance against leaf miner *Leucoptera coffeella* expressed by genetically transformed *Coffea canephora* in a pluriannual field experiment in French Guiana. *Euphytica*, 144: 321-329.

**Leroy T**, Marraccini P, Dufour M, Montagnon C, Lashermes P, Sabau X, Ferreira LP, Jourdan I, Pot D, Andrade AC, Glaszmann JC, Vieira LGE, Piffanelli P (2005). Construction and characterization of a *Coffea canephora* BAC library to study the

organization of sucrose biosynthesis genes. *Theoretical and Applied Genetics*, 111: 1032-1041.

Montagnon C, **Leroy T**, Cilas C, Charrier A (2003). Heritability of *Coffea canephora* yield estimated from several mating designs. *Euphytica*, 133: 209-218.

Montagnon C, **Leroy T**, Bertrand B, Charmetant P, Dufour M (2002). Résultats récents pour l'amélioration génétique du caféier. *Recherche et caféiculture, Plantations, recherche, développement*, ISSN 1254-7670: 85-94.

**Leroy T**, Henry AM, Royer M, Altosaar I, Frutos R, Duris D, Philippe R (2000). Genetically modified coffee plants expressing the *Bacillus thuringiensis* cry1Ac gene for resistance to leaf miner. *Plant Cell Reports*, 19 (4): 382-389.

Montagnon C, Cilas C, **Leroy T**, Yapo A, Charmetant P (2000). Genotype-location interactions for *Coffea canephora* yield in the Ivory Coast. *Agronomie*, 20: 101-109.

Montagnon C, Guyot B, Cilas C, **Leroy T** (1998). Genetic parameters of several biochemical compounds from green coffee, *Coffea canephora*. *Plant Breeding*, 117: 576-578.

Montagnon C, **Leroy T**, Eskes AB (1998). Amélioration variétale de *Coffea canephora*. I. Critères et méthodes de sélection. *Plantations, recherche, développement*, 5(1): 18-33.

Montagnon C, **Leroy T**, Eskes AB (1998). Amélioration variétale de *Coffea canephora*. II. Les programmes de sélection et leurs résultats. *Plantations, recherche, développement*, 5(2): 18-31.

**Leroy T**, Montagnon C, Cilas C, Yapo A, Charmetant P, Eskes AB (1997). Reciprocal recurrent selection applied to *Coffea canephora* Pierre. III. Genetic gains and results of first cycle intergroup crosses. *Euphytica*, 95: 347-354.

Baudouin L, Baril C, Clement Demange A, **Leroy T**, Paulin D (1997). Recurrent selection of tropical tree crops. EUCARPIA, Meeting on tropical plants; 1996/03/11-15; Montpellier (FRA). *Euphytica*, 96: 101-114.

Moschetto D, Montagnon C, Guyot B, Perriot JJ, **Leroy T**, Eskes AB (1996). Studies on the effect of genotype on cup quality of *Coffea canephora*. *Tropical Science*, 36: 18-31.

**Leroy T**, Montagnon C, Cilas C, Charrier A, Eskes AB (1994). Reciprocal recurrent selection applied to *Coffea canephora* Pierre. II. Evaluation of genetic parameters. *Euphytica*, 74 (1-2): 121-128.

Montagnon C, **Leroy T**, Kebe I, Eskes AB (1994). Importance de la rouille orangée et facteurs impliqués dans l'évaluation de la résistance au champ de *Coffea canephora* en Côte d'Ivoire. *Café, Cacao, Thé*, 38 (2): 103-112.

Montagnon C, **Leroy T**, Cilas C, Eskes AB (1993). Differences among clones of *Coffea canephora* in resistance to the scolytid coffee-twig borer. *International Journal of Pest management*, 39 (2): 204-209.

Montagnon C, **Leroy T**, Yapo A (1993). Caractérisation et évaluation de caféiers *Coffea canephora* prospectés dans des plantations de Côte d'Ivoire. *Café, Cacao, Thé*, 37(2): 115-119.

**Leroy T**, Montagnon C, Charrier A, Eskes AB (1993). Reciprocal recurrent selection applied to *Coffea canephora* Pierre in Côte d'Ivoire. I. Characterization and evaluation of breeding populations and value of intergroup hybrids. *Euphytica*, 67 (1): 113-125.

Montagnon C, **Leroy T** (1993). Réaction à la sécheresse de jeunes caféiers *Coffea canephora* de Côte d'Ivoire appartenant à différents groupes génétiques. *Café, Cacao, Thé*, 37 (3): 179-190.

Montagnon C, **Leroy T**, Yapo A (1992). Diversité génotypique et phénotypique de quelques groupes de caféiers (*Coffea canephora* Pierre) en collection. Conséquences sur leur utilisation en sélection. *Café, Cacao, Thé*, 36 (3): 187-198.

**Leroy T**, Charmetant P, Yapo A (1991). Application de la sélection récurrente réciproque au caféier *Coffea canephora* Pierre: premiers résultats du programme réalisé en Côte d'Ivoire. *Café, Cacao, Thé*, 35: 95-103.

Charmetant P, **Leroy T**, Bontems S, Delsol E (1990). Evaluation d'hybrides de *Coffea canephora* produits en champs semenciers en Côte d'Ivoire. *Café, Cacao, Thé*, 34 (4): 257-264.

### **Chapitres d'ouvrage**

**Leroy T**, Alpizar E, Dufour M, Etienne H (2006) Coffee (*Coffea* sp.): In *Agrobacterium* protocols, Second edition: volume II (Methods in molecular biology, vol.44), K Wang (ed.), pp 191-208. Humana press, Totowa, NJ, USA.

**Leroy T**; Dufour M (2004). *Coffea spp.* Genetic transformation. In *Transgenic crops of the world- Essential protocols*, IS Curtis (ed.), pp 159-170. Kluwer Academic Publishers, London, Great Britain.

Eskes AB, **Leroy T** (2004). Coffee selection and breeding. In *Coffee: growing, processing, sustainable production*, JN Wintgens (ed.), pp 57-86. Wiley-VCH Weinheim, Germany.

Montagnon C, **Leroy T**, Onzima R, Dufour M (2003). Pourquoi pas des terroirs Robusta? In *Cafés: terroirs et qualité*, C Montagnon (ed.), pp 145-152. CIRAD-MOCA, Montpellier, France.

Dufour M, **Leroy T**, Carasco-Lacombe C, Philippe R, Fenouillet C (2000). Coffee (*Coffea* sp.) genetic transformation for insect resistance. In *Coffee biotechnology and quality*; T Sera, CCR Soccol, A Pandey, S Roussos (Eds), pp 209-217. Kluwer, Dordrecht, The Netherlands.

Spiral J, **Leroy T**, Paillard M, Petiard V (1999). Transgenic Coffee (*Coffea* species). In *Biotechnology in Agriculture and Forestry, Vol. 44 Transgenic trees*, YPS Bajaj (ed.), pp 55-76. Springer-Verlag Berlin Heidelberg, Germany.

### **Mémoires et documents divers**

Leroy Thierry, 1998. Demande d'expérimentation au champ pour des plantes de *Coffea canephora* génétiquement modifiées pour la résistance à la mineuse des feuilles. Cirad-CP, présentation devant la commission du génie biomoléculaire (CGB) le 19 août 1998, 28 pages et Annexes.

Leroy Thierry, 1993. Diversité, paramètres génétiques et amélioration par sélection récurrente réciproque du caféier *Coffea canephora* P. Thèse de Doctorat de l'ENSA de Rennes, 147 pages.

## IV. SYNTHÈSE DES RESULTATS SCIENTIFIQUES

### *L'amélioration génétique des caféiers et de la qualité des cafés, du champ aux gènes*

#### **Introduction**

Le café est en valeur un des produits les plus importants dans les échanges internationaux. Deux espèces du genre *Coffea* ont une importance économique : *Coffea arabica*, cultivé dans les zones intertropicales d'altitude (500 à 2000 mètres) et *Coffea canephora* cultivé dans les zones de basse altitude, essentiellement en Afrique, au Brésil, en Indonésie et au Vietnam. Le café Arabica est réputé pour son arôme, et sa teneur en caféine modérée, alors que le Robusta, café produit par *C. canephora*, donne une boisson plus corsée et plus riche en caféine. Bien que reconnu comme de qualité moindre, le Robusta garde une part de marché de l'ordre de 35% car il convient à certains marchés, possède un meilleur potentiel extractible que l'Arabica, et est utilisé en mélange avec ce dernier en raison de son prix plus bas.

*Coffea arabica*, espèce tétraploïde autogame ( $2n=4x=44$ ), présente une grande diversité phénotypique (Charrier, 1978), mais la diversité génétique est faible (Anthony et al. 2001). L'amélioration de l'Arabica est basée sur deux groupes (Typica et Bourbon) dans lesquels des lignées ont été sélectionnées par sélection généalogique. Un mutant nain appelé Caturra a été découvert dans une population de Bourbon, et constitue, avec les variétés Mundo Novo et Catuai, les variétés les plus couramment cultivées. Le café produit par l'Arabica est considéré comme de bonne qualité, mais ce caféier souffre de sa sensibilité aux maladies, en particulier la rouille orangée et les nématodes. C'est pourquoi des programmes d'introggression à partir de *C. canephora* ont été menés pour apporter à l'Arabica des résistances, en créant les variétés Catimor, issues de l'hybridation des Caturra avec des hybrides interspécifiques (*C. arabica* x *C. canephora*) résistants aux pathogènes. L'amélioration de *C. arabica* porte donc essentiellement sur la productivité et la résistance aux maladies et insectes. La qualité du café produit est un critère de sélection important, en particulier dans les variétés introgressées, afin de maintenir un niveau élevé qui assure au planteur des prix attractifs (Bertrand et al. 2003).

L'amélioration de *C. canephora* a débuté au début du 20<sup>ème</sup> siècle à Java (Cramer, 1957). Des efforts de sélection se sont poursuivis dans plusieurs pays, en particulier en Afrique. La sélection de ce caféier diploïde et allogame a longtemps reposé sur une sélection massale dans des descendances contrôlées ou libres, avec une sélection de têtes de clones pour une multiplication végétative. Au début des années 1980, un programme de sélection récurrente réciproque basé sur la structure de la diversité génétique de l'espèce a été initié. Il utilise les populations correspondant aux deux groupes génétiques identifiés par Berthaud (1986) : le groupe Congolais d'Afrique centrale et le groupe Guinéen d'Afrique occidentale. Ce programme de sélection, outre les aspects de productivité des arbres, avait pour but l'amélioration des qualités technologiques et organoleptiques du café produit, le café Robusta souffrant dans ce domaine de la comparaison avec les Arabicas.

La démarche suivie tout au long des travaux présentés dans le présent document vise à l'amélioration des caféiers et de la qualité des cafés. Cette amélioration, toujours entreprise en collaboration avec les pays producteurs, qu'ils soient d'Afrique, d'Asie ou d'Amérique, vise à une meilleure productivité et une meilleure qualité des produits. La tolérance aux stress biotiques et abiotiques est par ailleurs indispensable pour une bonne adaptation des plantes dans chaque environnement.

Depuis plus de 20 ans, mes activités de recherche en amélioration des caféiers et de la qualité des cafés produits se sont déroulées principalement sur *C. canephora* pour la partie initiale, menée en Côte d'Ivoire (RCI), avant de s'élargir à *C. arabica* dans une deuxième phase, du transfert de gènes à la génomique de la qualité. Les nouvelles technologies, disponibles depuis quelques années sur les plantes modèles, ont été intégrées progressivement dans les travaux réalisés. L'ensemble des travaux entrepris s'est fait dans le cadre de projets impliquant des collaborations internationales (Côte d'Ivoire, Ouganda, Brésil, Inde), et nationales (Nestlé et IRD), avec une constante qui est l'animation d'une équipe aux contours et cadres de travail très variés dans les différentes étapes professionnelles. Les contraintes liées aux partenariats et aux situations politiques ou économiques dans les pays producteurs ont souvent posé les limites de nos travaux. Ainsi, les collaborations avec la RCI ont connu des contours très variables en fonction de la situation politique dans ce pays.

Les travaux présentés vont de la plante entière au champ jusqu'aux gènes d'intérêt en intégrant les méthodologies adaptées pour l'analyse du génome. Le travail sera présenté en quatre parties :

- A- Analyse de la diversité,
- B- Héritabilité des caractères, gains génétiques et index de sélection,
- C- Amélioration des caféiers par transfert de gènes,
- D- Génomique de la qualité des cafés : analyses de la maturation, gènes impliqués et leur expression.

Le projet de recherche présenté découle naturellement des travaux présentés et en cours. Il se concentre sur les possibilités d'études d'association chez *C. canephora*, et analyse les contraintes et la méthodologie pour mettre en place de telles études, et ainsi initier une sélection assistée par marqueurs des caféiers.

## **A Analyse de la diversité**

Les premières appellations en terme de diversité portaient sur des types phénotypiques décrits par plusieurs auteurs (Thomas, 1940 ; Chevalier, 1947 ; Cordier, 1961). Parmi les principaux types définis chez *C. canephora*, on peut retenir les « Robusta », les « Kouilou » d'Afrique, les « Conilon » du Brésil, les « Nganda » et les « Erect » d'Ouganda.

### **1 Evaluation de la diversité de *C. canephora* en Côte d'Ivoire**

Lorsque nos travaux ont débuté en Côte d'Ivoire en 1984, Berthaud avait mis en évidence une première structure génétique basée sur les marqueurs isoenzymatiques (Berthaud, 1986). La classification en Guinéens et Congolais reposait aussi sur des données biogéographiques et paléo climatiques définissant des zones refuges lors des dernières glaciations dans ces régions

(White, 1979). La définition des groupes Guinéens et Congolais par Berthaud s'est accompagnée de l'observation d'une forte productivité des arbres définis comme hybrides entre les deux groupes Guinéens et Congolais. Cette observation a servi de base au programme de sélection récurrente dont j'ai assuré la mise en place et le suivi pendant neuf ans en Côte d'Ivoire.

Pour mettre en place ce nouveau programme d'amélioration, je disposais d'une équipe d'une dizaine de techniciens expérimentés qui connaissaient très bien les caféiers au champ, et je travaillais avec un jeune chercheur ivoirien et le soutien d'un chercheur plus expérimenté basé à Abidjan. L'équipe travaillait aussi en collaboration avec des collègues de l'ORSTOM (IRD), plus spécialisés sur les caféiers sauvages. Sur la station de Divo, dans le centre ouest de la RCI, nous disposions de collections de caféiers cultivés (de l'ordre de 600 génotypes) et de collections de caféiers sauvages (près de 200 au départ de nos travaux).

La mise en place s'est faite selon plusieurs axes :

- élargissement de la base génétique de travail par des nouvelles prospections de caféiers sauvages et cultivés,
- évaluation de la diversité phénotypique et génotypique des collections de caféiers existantes,
- mise en place des plans de croisement pour le premier cycle de sélection récurrente réciproque.

#### *1-1 Elargissement de la base génétique*

Le caféier *C. canephora* a son aire naturelle de répartition qui couvre l'ensemble des zones forestières de Côte d'Ivoire. L'histoire de la culture dans ce pays (Portères, 1937 et 1959) montre qu'il y a eu des introductions d'Afrique centrale, en particulier de la zone Congolaise, et que ces caféiers introduits ont été cultivés en mélange avec les caféiers natifs Guinéens, sortis de la forêt par les planteurs. Ce sont ces mélanges de génotypes dans les plantations et la sélection effectuée par les planteurs qui ont permis la création d'hybrides entre les deux groupes qui se sont avérés productifs.

Lors du démarrage des opérations de sélection du caféier en Côte d'Ivoire dans les années 1960, les collections ont été mises en place avec deux sources d'approvisionnement principal : les caféiers sélectionnés introduits d'autres centres de recherche, en particulier de ceux localisés au Congo belge, et des caféiers prospectés localement dans les plantations de Côte d'Ivoire (Capot, 1977). Des collections de caféiers sauvages prospectés en RCA, au Cameroun et en Côte d'Ivoire complétaient ces collections de cultivés. Toutes ces collections ont servi de cadre d'étude à Berthaud (1986) pour les analyses isoenzymatiques.

L'élargissement de la base génétique, tant pour les caféiers cultivés en Côte d'Ivoire, que pour les caféiers sauvages de toute l'aire de répartition de *C. canephora* en Afrique de l'Ouest et centrale était d'un intérêt primordial pour le programme d'amélioration qui se mettait en place en 1984. C'est pourquoi, au cours des années 80 et 90, un certain nombre de prospections ont été effectuées :

- Caféiers sauvages prospectés au Congo en 1985 (De Namur et al. 1988) en Guinée en 1987 (Le Pierres et al. 1990), et en Côte d'Ivoire entre 1985 et 1991 (Le Pierres et al. 1990). L'ensemble de ces caféiers prospectés a été planté dans

les collections de la station de Divo, et analysé pour sa diversité génétique et phénotypique,

- Caféiers cultivés prospectés dans différentes régions de culture du centre et de l'ouest de Côte d'Ivoire. Ces prospections, effectuées dans des régions marginales pour la culture du caféier (pluviométrie inférieure à 1200mm), ont permis, avec l'aide des planteurs, de prélever des boutures des caféiers évalués comme bons producteurs par les planteurs, mais aussi d'essayer de sélectionner des génotypes Guinéens après une évaluation phénotypique par les prospecteurs.

Notre équipe a participé aux prospections de caféiers sauvages en Côte d'Ivoire et en Guinée avec les collègues de l'IRD, et a été maître d'œuvre pour les prospections en plantations. L'ensemble des caféiers prospectés a été ensuite planté dans les collections de Côte d'Ivoire. Il est remarquable de noter que la prospection dans les plantations dirigée vers le prélèvement de caféiers de type Guinéen a été une réussite puisqu'une vingtaine d'arbres de ce type a pu être prospectée, soit plus de 10% des arbres prospectés.

### *1-2 Evaluation de la diversité phénotypique et génotypique*

Les travaux d'évaluation avaient été débutés par Berthaud dans les laboratoires de l'ORSTOM à Adiopodoumé (Abidjan, 200 km de Divo). Ils ont été ensuite poursuivis par les collègues de l'ORSTOM pour les prospections de caféiers sauvages effectuées entre 1983 à 1985. Cependant, la collection de travail de génotypes introduits ou prospectés en plantation n'avait été génotypée qu'au tiers. Un laboratoire d'électrophorèse enzymatique a donc été construit par notre équipe sur la station de Divo. Opérationnel en 1987, il a permis d'étudier la diversité génotypique de l'ensemble des génotypes prospectés et introduits. Ces travaux ont été développés par C. Montagnon lors de son DAA en 1989 et ont fait l'objet de deux publications (Montagnon et al. 1992, 1993a).

L'évaluation génotypique a permis de préciser la structure de l'espèce *C. canephora*, puisque le groupe Congolais a été subdivisé en deux sous-groupes (Montagnon et al. 1992), le sous-groupe SG1 qui regroupe des génotypes de l'Afrique centrale atlantique et le sous-groupe SG2 qui regroupe les génotypes d'Afrique centrale continentale (RDC, Sud de la RCA, Sud-est Cameroun). Les résultats de ces travaux sont présentés sur la Figure 1. Cette structuration a été encore affinée avec l'aide des marqueurs RFLP qui ont permis de mettre en évidence l'existence de deux sous-groupes supplémentaires, B et C, dans le grand groupe Congolais (Dussert et al. 1999). La carte 1 présente la diversité génétique de *C. canephora* en Afrique telle qu'elle est connue en 2007 (Musoli, 2007).

Un certain nombre de paramètres génétiques ont été estimés sur les populations définies à partir des données RFLP (Montagnon, 2000). Les résultats montrent la forte structuration de l'espèce, avec une différenciation forte des Guinéens par rapport à toutes les populations Congolaises, elles-mêmes différenciées de façon équivalente les unes des autres. Les études récentes menées dans le cadre des thèses de P. Cubry et P. Musoli avec des marqueurs microsatellites ont confirmé la structure de l'espèce avec les Guinéens et 4 sous-groupes de congolais (Cubry et al. 2005, 2007). Les résultats les plus récents montreraient une structuration du groupe Guinéen, une population prospectée en Côte d'Ivoire dans une région sèche du centre du pays (population Pélézi, voir Figure 1) présente en effet des caractéristiques différentes des autres populations guinéennes (Cubry, comm. pers.).



Tableau1- Caractéristiques phénotypiques des populations Guinéennes et Congolaises de *C. canephora*. (d'après Montagnon, 2000).

		CONGOLAIS			GUINEEN	
		SG1	SG2			
	Caractère	Cultivés	Cultivés	Sauvages	Cultivés	Sauvages
	Effectif	84	105	36 à 90	39	70 à 135
Feuilles	Longueur des feuilles (mm)	188	203	221	170	176
	Largeur des feuilles (mm)	71	82	93	68	75
	Forme des feuilles (Longueur/largeur)	2,68	2,49	2,38	2,53	2,36
	Acumen	Court à long	Long	Long	Long	Long
	Pétiole	Court à long	Moyen à long	Long	Court à long	Moyen à long
	Domaties	Absentes à bien visibles	Peu marquées à bien visibles	Peu marquées à bien visibles	Peu marquées à bien visibles	Peu marquées à bien visibles
	Fruits	Pédoncule fruit	Court à moyen	Court à moyen	Court à long	Court à long
Taille du disque		Petit	Petit	Petit	Grand	Petit
Relief du disque		Plat à saillant	Saillant	Plat à saillant	Plat	Plat
Granulométrie (g. pour 100 grains 12% humidité)		14,2	15,4	9,5	10,0	9,0
Divers	Ramification	Moyenne à forte	Faible à moyenne	Faible à moyenne	Moyenne à forte	Moyenne à forte
	Longueur Entrenoeuds	Moyen à court	Long	Long	Court	Court
	Phénologie	Très tardifs	Moyennement tardif	Moyennement tardif	Précoces	Précoces
	Caféine (% matière sèche)	2,7	2,3	2,3	2,9	2,6
	Rouille orangée (% de plants résistants)	73	84	97	31	28
	Scolyte des branchettes (% de plants résistants)	94	73	50	88	77
	Résistance à la sécheresse	Bonne	Faible	Faible	Bonne	Bonne
	Qualité organoleptique	Bons arômes et acidité Amertume et corps moyens	Bons arômes et acidité Faibles amertume et corps		Faibles arômes et acidité Fortes amertumes et corps	

A côté de cette évaluation de la diversité génétique, une évaluation phénotypique a été entreprise dans l'ensemble des collections. Deux stagiaires de deuxième année de l'ISTOM

ont participé à cette évaluation (C. Duchemin et F. Bodin). Cette évaluation a concerné les aspects suivants :

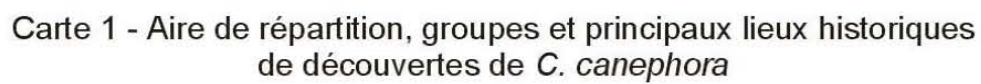
- description botanique des génotypes : longueur des entre-nœuds, forme des feuilles,
- notation des aléas biotiques et abiotiques,
- caractéristiques du café produit : taille des grains, taux de fèves caracoli, taux de caféine, rendement en café marchand,
- productivité des génotypes,
- croisements réalisés avec les génotypes en collection.

Une base de données sous Foxpro a été créée pour cette collection afin de regrouper ces données avec l'évaluation par des marqueurs isoenzymatiques, et les informations sur les croisements réalisés et les essais mis en place avec ces génotypes. Cette base de données, construite par l'équipe avec l'aide de VSN (Volontaires du Service National) basés sur la station, a été finalisée par Louis Marie Raboin en 1994 dans le cadre d'un certificat de spécialisation informatique.

Les caractéristiques générales des génotypes « Guinéens » et « Congolais » sont présentées dans le Tableau 1, pour les Guinéens, les congolais SG1 et SG2. Les groupes B et C présentent aussi des caractéristiques intéressantes. En particulier, les génotypes du groupe C (caféiers de la Nana) présentent une architecture particulière de type « nain » très ramifiée avec des entre-nœuds courts.

Le renforcement de l'équipe avec l'arrivée de C. Montagnon en 1989 a permis d'amplifier les travaux sur la diversité des caféiers dans les collections et les premiers essais en cours. Des études phénotypiques précises ont été menées pour certains aléas biotiques et abiotiques d'importance en Côte d'Ivoire. Des travaux ont ainsi été faits sur la réaction à la sécheresse des caféiers Guinéens, Congolais et hybrides (Montagnon et Leroy, 1993), sur l'importance de la rouille orangée *Hemileia vastatrix* et l'évaluation de sa résistance au champ (Montagnon et al. 1994) et sur les différences clonales dans la résistance au scolyte des branchettes *Xylosandrus compactus* (Montagnon et al. 1993b).

Enfin, la qualité organoleptique des cafés a fait l'objet d'analyses comparées entre les cafés produits par les Guinéens, les Congolais et les hybrides. Les Congolais sont plus appréciés que les Guinéens, et les hybrides montrent une grande variabilité (Moschetto et al. 1996). La qualité organoleptique des cafés produits, mais aussi leur qualité technologique (poids de 100 fèves) et industrielle (extractibilité) font partie des critères de sélections dans les sorties variétales de la sélection récurrente. La liste des critères de sélection et leurs objectifs est présentée dans le Tableau 2 (Montagnon, 2000).



- 19

### *1-3 Mise en place du programme de sélection récurrente réciproque*

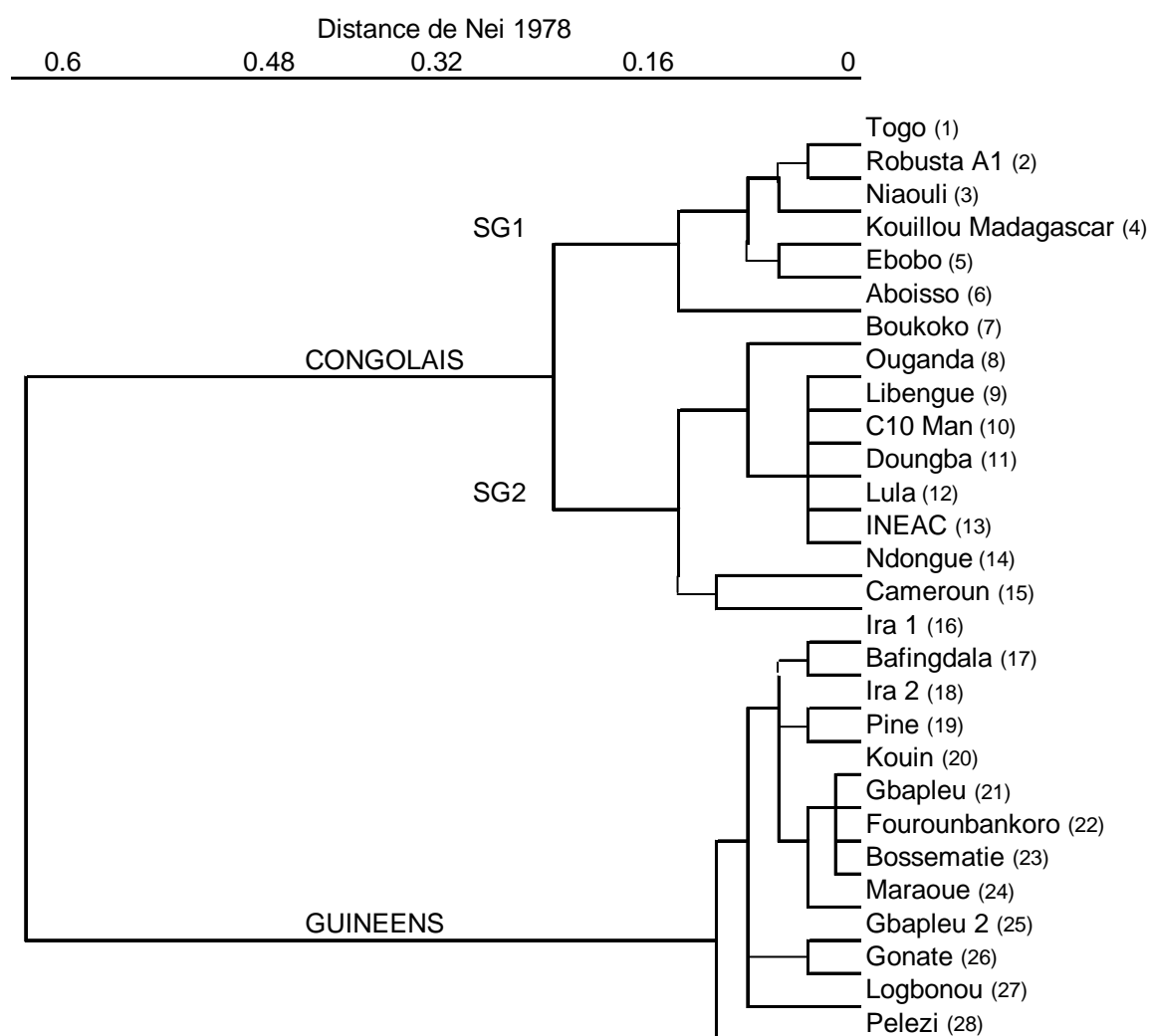
L'amélioration par sélection récurrente réciproque du caféier *C. canephora* a été mise en place à partir de 1984 suite aux travaux de Berthaud sur la structuration de l'espèce (Guinéens et Congolais) et la valeur des hybrides naturels entre ces deux groupes prospectés dans les plantations de Côte d'Ivoire. Les caractéristiques complémentaires des deux groupes de diversité sont aussi un élément important pour la mise en place de ce programme : résistance à la sécheresse, ramification importante, petite taille des arbres et précocité pour les Guinéens, faible sensibilité à la rouille orangée, taux de caféine plus faible, forte granulométrie et tardiveté de la récolte pour les Congolais.

Le schéma général de sélection est présenté sur la Figure 2. Le modèle adopté sur le caféier était la sélection du palmier à huile, plante sur laquelle une telle sélection était en place depuis plusieurs années, avec des caractéristiques de l'espèce (pérennité, allogamie, groupes différenciés) proches de celles du caféier (Meunier et Gascon, 1972). Des contacts ont d'ailleurs été noués dès 1985 avec les sélectionneurs travaillant sur ce même type de sélection, au Cirad et dans les centres INRA de Bordeaux, Lusignan et Montpellier.

Les données préliminaires connues sur la transmission des caractères dans l'espèce (Bouharmont et al. 1986) et la variabilité phénotypique des caféiers en collection (Leroy et Charmetant, 1987) ont permis de faire le choix d'une cinquantaine de génotypes de chacun des deux groupes Guinéens et Congolais. Compte tenu de la prédominance des aptitudes générales à la combinaison sur cette espèce, la sélection indirecte par testeurs de la population complémentaire a été préférée à une sélection directe. Les testeurs (trois pour chaque population) ont été choisis à partir de travaux antérieurs (Capot, 1977).

Les premiers résultats de la mise en place de ces populations et de la valeur des hybrides intergroupes ont été publiés dans deux articles (Leroy et al. 1991, 1993). La valeur des hybrides intergroupes est confirmée par ces premiers résultats, des hybrides aussi productifs et aussi homogènes génétiquement que les clones vulgarisés ont été obtenus dès les premières années dans ce programme d'amélioration. La randomisation totale arbre par arbre avait été choisie pour les sols très hétérogènes de Côte d'Ivoire. La faible qualité et l'hétérogénéité des sols tropicaux de la station de Divo nous a amené à mettre en place des dispositifs d'évaluation adaptés : plus de 50 pieds analysés par descendance, dispositifs de bordures hors essais, analyse des résultats en utilisant les données des plus proches voisins (analyses de covariance de Papadakis). Outre l'amélioration des populations, la sortie variétale a été développée sous forme de clones (sélection des meilleurs arbres dans les essais), mais aussi d'hybrides de clones, en raison de leur forte valeur agronomique et de leur homogénéité génétique.

Un système d'essais multi locaux d'hybrides et de clones a aussi été mis en place, en tenant compte des interactions génotype x environnement (Montagnon et al. 2000). Ces essais, au nombre d'une quinzaine, couvraient l'ensemble des conditions édapho-climatiques de la culture du caféier en Côte d'Ivoire.



#### Description des origines étudiées

Pays	Origine sylvestre	Origine cultivée
Cameroun	15	
Rép. Centrafricaine	9;11;14	7
Côte d'Ivoire	16-18;20-28	
Guinée	19	
R. D. Congo		10;12;13
Ouganda		8
Gabon + Sud Congo		1, 2, 3, 4
Inconnu		5, 6

Figure 1 - Dendrogramme illustratif de la diversité génétique du matériel végétal de *Coffea canephora* disponible pour la sélection obtenu à partir de données iso-enzymatiques sur 9 loci (d'après Montagnon *et al*, 1992b)

Le premier cycle de sélection a duré environ une quinzaine d'années entre les premiers croisements intergroupes du premier cycle et ceux du deuxième cycle. Plusieurs centaines de croisements intergroupes, mais aussi intragroupes ont été réalisés chaque année. L'équipe a donc mis en place des protocoles standard pour un certain nombre d'activités : pollinisations, germinations et pépinières, notations au champ, saisie des données, analyses. L'arrivée des premiers ordinateurs en 1984 a permis de mettre au point une gestion rationnelle des essais et des données. Des programmes informatiques spécifiques pour nos essais de grande taille en randomisation totale avec des effectifs inégaux ont été développés par l'équipe à partir de données saisies avec le logiciel « Statitcf ». Les algorithmes ont été développés par les chercheurs de l'équipe, des VSN ont ensuite effectué la programmation sous GWBASIC. Il était ainsi possible de suivre les essais depuis le tirage aléatoire des emplacements dans les parcelles, les données au jour le jour, leur analyse avec les transformations de variables éventuelles et les analyses de covariance, et enfin le repérage des arbres sur leur valeur propre.

## **2 Evaluation de la diversité de *C. canephora* en Ouganda**

Depuis 2003, l'équipe a été impliquée dans des travaux avec l'Ouganda. Cette région d'Afrique avait constitué dans les années 60 et 70 un centre de sélection important (Millot, 1974) pour *C. canephora*. La présence de caféiers sauvages dans ce pays était rapportée, mais ces génotypes n'avaient jamais fait l'objet de prospections et d'évaluations systématiques. Deux projets impliquant le Cirad ont débuté en 2003-2004. L'un portait sur la mise en évidence éventuelle de marqueurs de la qualité chez les caféiers cultivés en Ouganda, l'autre portait sur la recherche de sources de résistance à la trachéomycose *Fusarium xylarioides*, maladie vasculaire qui fait des ravages en Ouganda. Le premier projet, financé par l'USDA, impliquait une chercheuse ougandaise, Pauline Aluka, qui a effectué deux stages de 6 mois dans notre équipe pour analyser la diversité génétique de ses échantillons. Le deuxième projet, financé par l'Europe, impliquait un chercheur ougandais, Pascal Musoli, qui a réalisé sa thèse de doctorat dans le cadre de ce projet (Musoli, 2007). Outre la partie phytopathologie, réalisée dans un autre laboratoire du Cirad, l'ensemble des travaux sur la diversité génétique des caféiers sauvages a été réalisé dans notre équipe. Ces deux projets ont permis d'initier une coopération avec nos collègues ougandais par l'encadrement de ces deux chercheurs pendant trois ans.

Ces deux projets ont permis d'analyser la diversité génétique des caféiers d'Ouganda, tant sauvages que cultivés. Nous disposons de documents assez anciens (Thomas, 1940, 1944) décrivant la diversité et l'histoire des caféiers en Ouganda. Ces documents ont permis de retracer les débuts de la culture des caféiers en Ouganda, avec les deux « variétés » les plus connues, les « Erect » et les « Nganda », qui se différencient très bien par leurs caractéristiques phénotypiques. A partir des documents disponibles sur les collections de matériel vivant (collections en station, jardin botanique d'Entebbe) et les forêts susceptibles d'héberger des caféiers sauvages (forêts résiduelles de l'ouest du pays, Kibale et Itwara), un travail important d'échantillonnage a été réalisé :

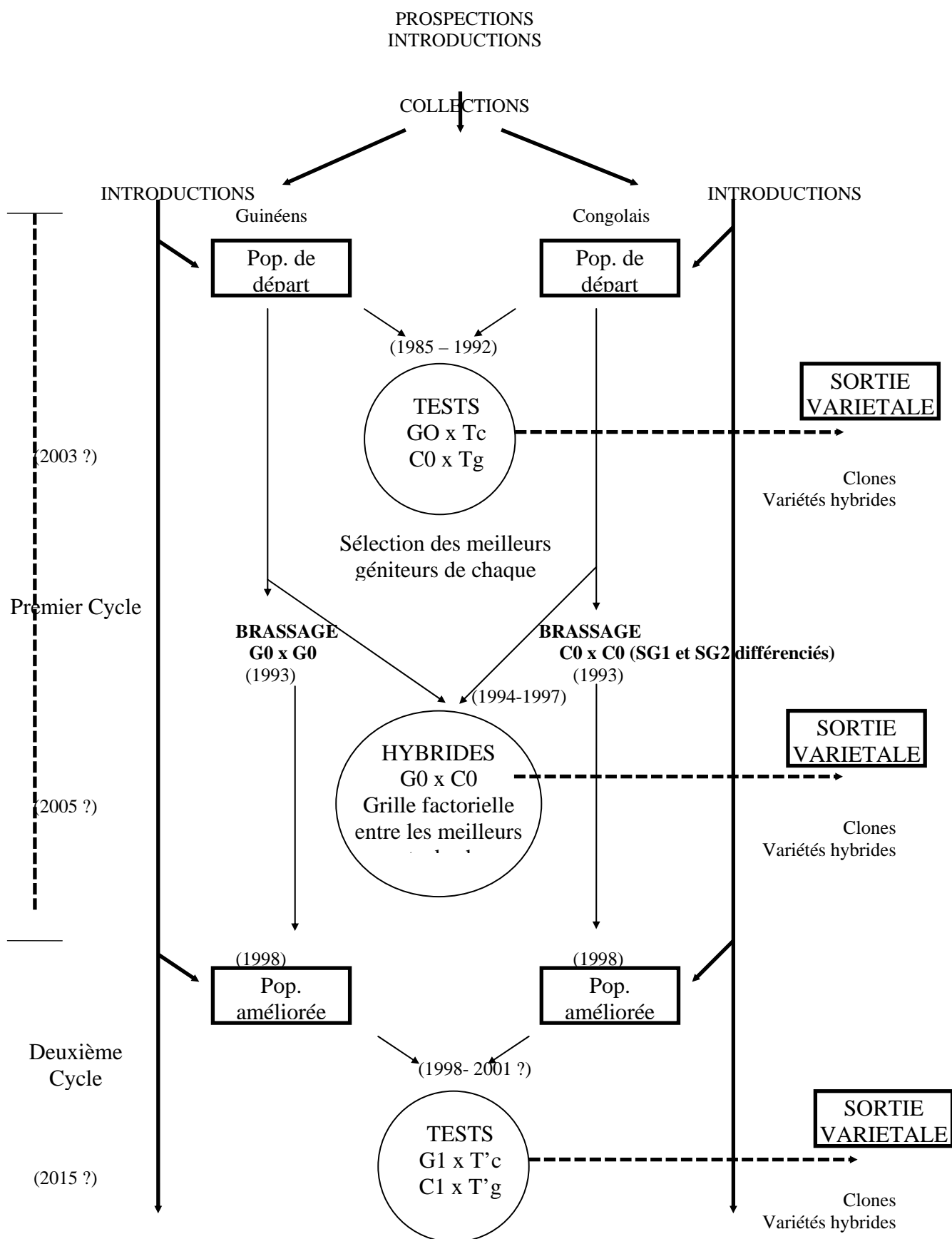


Figure 2 - Représentation schématique de la Sélection Récurrence Réciproque appliquée à *Coffea canephora* en Côte d'Ivoire (d'après Leroy, 1993).

- plus de 200 génotypes spontanés des forêts de Kibale et Itwara ont été récoltés dans une dizaine de lieux de collecte et analysés pour leur diversité génétique,
- des génotypes ont été échantillonnés dans le jardin botanique d'Entebbe, près de Kampala,
- des génotypes « Erect » et « Nganda » des collections de travail du centre de Recherches sur le café ont également été échantillonnés,
- un échantillonnage systématique de caféiers dans les plantations d'une dizaine de districts du pays a été effectué,
- des arbres anciennement cultivés (matériel féral) ont été récoltés sur des îles du lac Victoria.

Tous les arbres échantillonnés ont été analysés en diversité avec des marqueurs microsatellites, et des analyses globales ont été menées sur le matériel d'Ouganda en relation avec les groupes génétiques connus (voir Figure 3). Ces travaux ont permis de mesurer un certain nombre de paramètres de génétique des populations : différenciation entre les populations, isolation par la distance dans les populations sauvages, structure des populations et relations entre celles-ci. Les résultats montrent l'originalité du matériel sauvage d'Ouganda par rapport au matériel déjà connu. L'Ouganda pourrait donc être considéré comme un centre secondaire de diversité de l'espèce. Il n'est pas exclu que des zones forestières refuges aient pu exister dans la zone actuelle de forêts résiduelles, et que les caféiers aient pu y survivre lors des dernières glaciations (Maley, 1996 ; Jolly et al. 1997).

D'autre part, l'analyse des phénotypes « Erect », « Nganda » et des caféiers d'origine férale a montré que ces plantes étaient très proches des caféiers cultivés du sous-groupe SG2 dans la cuvette congolaise. Des échanges de matériel végétal au début du 20<sup>ème</sup> siècle, et les transferts de matériel d'une collection à l'autre par voie de graines peuvent expliquer ces forts apparentements génétiques, alors que les génotypes « Nganda » sont vraisemblablement au départ d'origine sauvage. Quant aux caféiers des îles, il s'avère que leur isolement géographique leur a permis d'évoluer de manière divergente depuis trois à cinq générations. Ces caféiers sont en effet génétiquement différents des autres groupes de caféiers cultivés et ils possèdent des caractéristiques intéressantes comme leur plus forte tolérance à la trachéomycose.

Ces analyses de matériel d'Ouganda présentent un intérêt très important pour l'amélioration des caféiers :

- des caféiers sauvages ont pu être prélevés dans des forêts et mis en collection pour une évaluation agronomique ultérieure,
- un nouveau groupe génétique a été mis en évidence, qui pourrait être intégré dans les programmes d'amélioration (sélection récurrente),
- des sources de tolérance à la trachéomycose ont été trouvées dans certaines populations comme les populations férales.

Les résultats sont en cours de publication (Musoli et al. article soumis).



Tableau 2 - Liste des critères de sélection en fonction des objectifs d'amélioration de *C. canephora* (Montagnon, 2000).

	Objectifs				
	Amélioration de la qualité du café produit	Réduction des intrants		Réduction de la main d'oeuvre	Préservation de l'environnement
Critère de sélection		Engrais Pesticides	Herbicides		
Augmentation taille des grains	+++			++	
Critères organoleptiques	+++				
Teneur en caféine	?				
Potentiel extractible	+				
Maturité groupée	+++	+		+++	+
Vigueur au jeune âge					++
Optimum de production à faible dose d'engrais		+++		+	++
Résistance/tolérance rouille orangée		++		+	++
Résistance au scolyte des baies	++	+++		+	+++
Résistance au scolyte des branchettes		+			+
Architecture: Petite taille, ramification			++	+++	+++
Adaptation à l'écimage			+++	+	
Faible production de gourmands				+++	

### 3 Diversité de *C. arabica* et autres espèces

Des travaux sur la diversité génétique de *C. arabica* ont montré que celle-ci était faible dans les génotypes cultivés comme sauvages (Anthony et Lashermes, 2005), avec cependant une séparation entre les génotypes sauvages de l'ouest et de l'est du rift en Ethiopie (Montagnon et Bouharmont, 1996).

Dans le cadre des travaux de l'équipe sur la diversité des caféiers, une analyse de la diversité du genre *Coffea* par des microsatellites a été menée. Cette étude était la première à porter sur un nombre important de marqueurs microsatellites (une quarantaine) et sur une palette d'espèces représentant l'ensemble de la diversité du genre. Elle a été menée dans le cadre de la thèse de P. Cubry a permis de confirmer la structuration par origine géographique qui avait déjà été mise en évidence avec d'autres marqueurs (Lashermes et al. 1997). Les différents origines (Madagascar, Afrique de l'Est, Afrique centrale et Afrique de l'ouest et centrale) se retrouvent, mais ce dernier groupe est séparé en deux : les espèces *arabica*, *congensis* *canephora* et *brevipes* sont groupées, alors que les espèces *liberica*, *humilis* et *stenophylla* sont positionnées dans un autre sous-groupe (Cubry et al. 2008). Ces travaux permettent d'envisager une meilleure utilisation des espèces sauvages du genre pour l'amélioration des deux espèces cultivées en utilisant au mieux la diversité existante, les caractéristiques de chaque espèce et les relations génétiques entre les taxons.

#### Principales publications de l'équipe

Cubry P, Musoli P, Legnaté H, Pot D, De Bellis F, Poncet V, Anthony F, Dufour M, **Leroy T** (2008). Diversity in coffee using SSR markers: structure of the *Coffea* genus and perspectives for breeding. *Genome*, 51: 50-63.

Cubry P, Musoli P, Legnaté H, Aluka P, Dufour M, Pot D, De Bellis F, Leroy T (2007). Genetic diversity within *Coffea canephora* germplasm maintained in RCI using microsatellites loci. Results and future prospects. In 21<sup>st</sup> International Conference on Coffee Science (ASIC), Montpellier, France, September 2006, 11-15, pp. 912-915. ASIC, Paris, France.

Cubry P, De Bellis F, Pot D, Musoli P, Legnaté H, **Leroy T**, Dufour M (2005). Genetic diversity analyses and linkage disequilibrium evaluation in some natural and cultivated populations of *Coffea canephora*. In 4<sup>th</sup> Plant genomics European meeting, Amsterdam, September 2005, 20<sup>th</sup>-23<sup>th</sup>. Poster.

Le Pierres D, Charmetant P, Yapo A, **Leroy T**, Couturon E, Bontems E et Tehe H (1990). Les caféiers sauvages de Côte d'Ivoire et de Guinée. Bilan des missions de prospection effectuées de 1984 à 1987. In 13<sup>th</sup> International Conference on coffee science (ASIC), Paipa, Colombia, August 2009, pp. 420-428. ASIC, Paris, France.

**Leroy T**, Charmetant P (1987). Les descripteurs de *Coffea canephora* Pierre. In 12<sup>th</sup> International Conference on Coffee Science (ASIC), Montreux, Switzerland, June-July 1987, 29-03, pp. 411-417. ASIC, Paris, France.

**Leroy T**, Montagnon C, Charrier A, Eskes AB (1993). Reciprocal recurrent selection applied to *Coffea canephora* Pierre in Côte d'Ivoire. I. Characterization and evaluation of breeding populations and value of intergroup hybrids. *Euphytica*, 67 (1): 113-125.

**Leroy T**, Charmetant P, Yapo A (1991). Application de la sélection récurrente réciproque au caféier *Coffea canephora* Pierre: premiers résultats du programme réalisé en Côte d'Ivoire. *Café, Cacao, Thé*, 35: 95-103.

Montagnon C (2000). Optimisation des gains génétiques dans le schéma de sélection récurrente réciproque de *Coffea canephora* Pierre. ENSA Montpellier, France, PhD thesis.

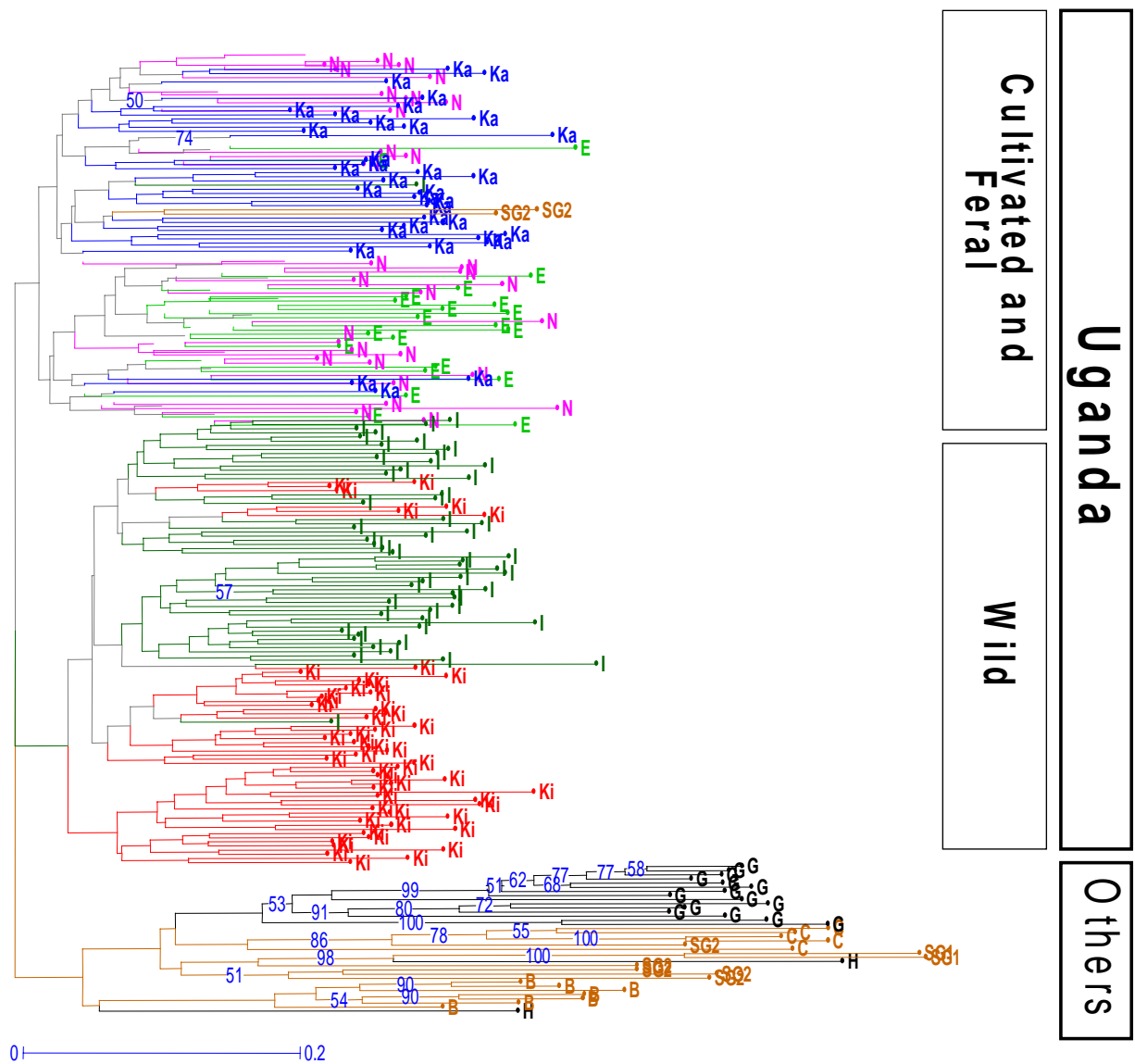


Figure 3- Diversité génétique de *C. canephora* en Ouganda  
En relation avec les autres groupes de diversité indentifiés

Montagnon C, Cilas C, **Leroy T**, Yapo A, Charmetant P (2000). Genotype-location interactions for *Coffea canephora* yield in the Ivory Coast. *Agronomie*, 20: 101-109.

Montagnon C, **Leroy T**, Kebe I, Eskes AB (1994). Importance de la rouille orangée et facteurs impliqués dans l'évaluation de la résistance au champ de *Coffea canephora* en Côte d'Ivoire. *Café, Cacao, Thé*, 38 (2): 103-112.

Montagnon C, **Leroy T** (1993). Réaction à la sécheresse de jeunes caféiers *Coffea canephora* de Côte d'Ivoire appartenant à différents groupes génétiques. *Café, Cacao, Thé*, 37 (3): 179-190.

Montagnon C, **Leroy T**, Yapo A (1993a). Caractérisation et évaluation de caféiers *Coffea canephora* prospectés dans des plantations de Côte d'Ivoire. *Café, Cacao, Thé*, 37(2): 115-119.

Montagnon C, **Leroy T**, Cilas C, Eskes AB (1993b). Differences among clones of *Coffea canephora* in resistance to the scolytid coffee-twig borer. *International Journal of Pest management*, 39 (2): 204-209.

Montagnon C, **Leroy T**, Yapo A (1992). Diversité génotypique et phénotypique de quelques groupes de caféiers (*Coffea canephora* Pierre) en collection. Conséquences sur leur utilisation en sélection. *Café, Cacao, Thé*, 36 (3): 187-198.

Moschetto D, Montagnon C, Guyot B, Perriot JJ, **Leroy T**, Eskes AB (1996). Studies on the effect of genotype on cup quality of *Coffea canephora*. *Tropical Science*, 36: 18-31.

Musoli P (2007). Recherche de sources de résistance à la trachéomycose du caféier *Coffea canephora* Pierre, due à *Fusarium xylarioides* Steyaert en Ouganda. Thèse de Doctorat, ENSA Montpellier (France).

Musoli P, Cubry P, Aluka P, Billot C, Dufour M, De Bellis F, Pot D, Bieysse D, Charrier A, **Leroy T**. Genetic differentiation of wild and cultivated populations: diversity of *Coffea canephora* in Uganda. *Soumis à Genome*

## **B Transmission des caractères et gains génétiques par sélection classique**

Dès 1984 en Côte d'Ivoire, avec la mise en place de la sélection récurrente réciproque sur le caféier *C. canephora*, des centaines de croisements ont été faits entre des génotypes Guinéens et Congolais. Pour des raisons pratiques et de limitation des surfaces plantées à 5 hectares par an (soit près de 10000 arbres en observation individuelle), nous avons choisi la sélection par testeurs de la population complémentaire pour connaître la valeur en croisement de nos génotypes. Trois testeurs congolais et trois testeurs guinéens ont été croisés avec l'ensemble des génotypes de la population complémentaire. Les nouveaux génotypes introduits ou prospectés, en forêt comme en plantation ont été intégrés au fur et à mesure dans le programme d'amélioration.

Sur l'ensemble des essais mis en place lors de 8 ans du premier cycle de sélection, des analyses de transmission des caractères sélectionnés ont été réalisés. Les résultats sur l'héritabilité des caractères ont fait l'objet de trois publications (Leroy et al. 1994 ; Montagnon et al. 1998a ; Montagnon et al. 2003). De façon générale, l'héritabilité au sens strict des caractères de vigueur est moyenne (0.2 à 0.5) : diamètre au collet, longueur des rameaux plagiotropes, diamètre de la canopée, nombre d'entre-nœuds et longueur moyenne d'un entre-nœud. Les héritabilités au sens large sont souvent très élevées (supérieures à 0.6). Ces résultats ont été obtenus par analyse de variance ou par régression parents descendants avec des résultats comparables, même si les régressions donnent des héritabilités un peu plus fortes. Les héritabilités calculées sur les parents Guinéens et Congolais ne donnent pas de résultats différents. La productivité des arbres pour un cycle de production (5 récoltes) a

également été analysée. L'héritabilité des deux premières récoltes (à 2 et 3 ans) est généralement très élevée (supérieure à 0.6) alors que l'héritabilité pour le rendement cumulé est plus faible (de l'ordre de 0.3) avec une héritabilité au sens large qui reste très élevée (supérieure à 0.7). Ces résultats ont été confirmés et précisés en 2003 sur un grand nombre d'essais (11). La taille des fèves (granulométrie) présente des héritabilités au sens strict moyennes de l'ordre de 0.2 à 0.3.

Une étude spécifique sur un essai (Montagnon et al. 1998a) a permis d'évaluer l'héritabilité de quelques composantes biochimiques. Les héritabilités observées sont fortes pour la teneur des fèves en lipides et en caféine, moyenne pour les acides chlorogéniques, la trigonelline et la teneur en sucres. Ces résultats cohérents avec des études précédentes (Charrier et Berthaud, 1975 ; Bouharmont et al. 1986) présentent un grand intérêt pour l'amélioration des caféiers. Il sera évidemment plus facile d'améliorer la teneur en lipides que celle en sucres, cette dernière étant très dépendante de l'environnement.

Nous avons également évalué les corrélations génétiques entre les caractères de production et de vigueur. Des corrélations génétiques significatives, mais différentes suivant les essais, ont été mises en évidence entre diamètre au collet, encombrement de l'arbre et productivité (Leroy et al. 1994). Des données plus précises sur un grand nombre d'essais présentées par Montagnon (2000) mettent en évidence des corrélations plus faibles, mais souvent significatives.

Les gains génétiques obtenus lors du premier cycle de sélection ont été estimés à partir des données d'héritabilité individuelle (pour la sélection de clones) ou familiale (pour la sélection d'hybrides), des variances des caractères et de l'intensité de sélection appliquée dans chaque essai (Leroy et al. 1997). La sélection des meilleurs arbres dans les essais peut être réalisée par une sélection par index ou par niveaux indépendants. Une sélection par index a été proposée entre le rendement, la vigueur au jeune âge et le diamètre de la canopée des arbres à 4 ans. Il s'avère que les gains génétiques obtenus sont très importants pour la productivité : 40 à 120% suivant les essais, le type de sélection (individus ou familles) et le pourcentage de sélection. Les gains génétiques sont modérés pour le diamètre au collet, et sont faibles ou négatifs pour l'encombrement des arbres, car notre sélection favorise les arbres à faible encombrement afin de limiter la concurrence entre voisins. Une sélection par niveau indépendant montre des gains génétiques faibles pour la granulométrie. Les analyses postérieures ont confirmé les gains génétiques importants obtenus pour la productivité, et plus faibles pour la taille des fèves sur l'ensemble du premier cycle de sélection récurrente (Montagnon, 2000).

Le premier cycle de sélection récurrente a duré une dizaine d'années de 1984 à 1994. La sélection sur testeurs a été efficace, puisque le premier cycle de sélection a permis la sélection de quelques dizaines de têtes de clones à tester en conditions multi locales, et d'une quinzaine d'hybrides plus productifs et aussi homogènes que le clone vulgarisé le plus productif. A l'issue des tests intergroupes de tous les génotypes (environ 70 par groupe) plusieurs types d'action ont été menées :

- Une grille de croisements intergroupes entre les meilleurs géniteurs du premier cycle a été réalisée. L'évolution des connaissances sur l'espèce, avec la mise en évidence en 1989 des sous-groupes congolais SG1 et SG2 a été prise en compte dans cette grille de croisements.
- Des croisements entre des génotypes des groupes SG1 et SG2 ont été réalisés.

Tableau 3- Grille factorielle des croisements intergroupes des meilleurs parents (sélectionnés précocement) du premier cycle de SRR et année de plantation des croisements réalisés. Les cases vides représentent des croisements non réalisés (Montagnon, 2000).

		GUINEENS										CONGOLAIS SG1				
		410	155	O2189	O2213	O2214	O2224	O2255	O2183	O2056	O2009	466	307	A03	O72	O77
CONGOLAIS	SG2	464	96	94			94		94		94		96	94	96	96
		178	94	97	94	94		96	96	96	94	94	96		97	
		444	94		94	94	94	96	94	94	94	94		94	94	94
		222	94	96							96	97			95	96
		392	94	94		94	94		94	94	94	94		94		
		O57	94	97			97		94	94	94	97		97	97	96
		111	94		94	97	97	97	97			94	96	97	97	
		O32	97	94		94		97	97					94		
	SG1	466	94	97	94	94		94	94		94					
		307			97											
		A03	94	94	94	94	97	94	94	97	94					
		O72	95	95	95	95	94	97	94	95	97	94				
		O77	97		94	96						94				

Au total, une centaine de croisements ont été réalisés, dont une vingtaine entre les meilleurs géniteurs des sous-groupes SG1 et SG2. Ces hybrides ont permis la sélection de descendances et de clones à haut potentiel de production (Tableau 3).

Après avoir réalisé l'ensemble des croisements intergroupes sur testeur, des croisements de brassage intragroupe ont été réalisés. Ces croisements étaient basés sur les résultats disponibles en 1991. Les géniteurs les plus performants ont été utilisés comme femelles et croisés avec un mélange pollinique de ces mêmes géniteurs.

Une sélection par index a ensuite été réalisée sur la productivité des hybrides issus de ces géniteurs (dans les croisements intergroupes et intragroupes), la sensibilité à la rouille et la taille des fèves. Un certain nombre de demi-frères issus du brassage intra groupe ont ainsi été sélectionnés pour débiter le deuxième cycle de sélection. D'autre part, en plus des demi-frères sélectionnés, un certain nombre de géniteurs nouveaux non utilisés auparavant (groupes B et C) seront introduits tel quels dans le deuxième cycle de sélection. La valeur d'un certain nombre de géniteurs a été confirmée quand tous les essais ont été analysés en 1998, d'autres ont été écartés en raison en particulier de leur valeur en test intragroupe pour l'amélioration des populations en sélection. Pour la mise en place du deuxième cycle de sélection, les 4 sous groupes congolais (SG1, SG2, B et C) ont été considérés séparément. L'introduction des génotypes d'Ouganda dans ce programme d'amélioration serait intéressante à envisager. Enfin, les résultats du premier cycle ont également montré la pertinence de l'utilisation des testeurs dans notre programme d'amélioration (Montagnon et al. 2008).

Ce programme d'amélioration, initié en 1984, que j'ai suivi directement jusqu'en 1994, puis qui a été suivi par des personnes de l'équipe jusqu'en 2002, a permis une forte augmentation de la productivité potentielle des sorties variétale. En particulier, il a permis de sélectionner des hybrides intergroupes mais aussi intragroupes (Congolais) très productifs, et qui offrent l'avantage de pouvoir être distribués par graines aux planteurs par rapport à des clones multipliés par bouturage horticole. Outre la productivité, la taille des fèves a fait l'objet d'une sélection lors des cycles de sélection. Lors des phases avancées de la sélection, le taux de caféine du café produit par les clones ou hybrides, et la qualité à la tasse ont été évaluées, et prises en compte pour la sélection finale. Deux publications de synthèse ont été faites sur les résultats de ce programme de sélection en Côte d'Ivoire (Montagnon et al. 1998b et 1998c).

## Principales publications de l'équipe

**Leroy T**, Montagnon C, Cilas C, Yapo A, Charmetant P, Eskes AB (1997). Reciprocal recurrent selection applied to *Coffea canephora* Pierre. III. Genetic gains and results of first cycle intergroup crosses. *Euphytica*, 95: 347-354.

**Leroy T**, Montagnon C, Cilas C, Charrier A, Eskes AB (1994). Reciprocal recurrent selection applied to *Coffea canephora* Pierre. II. Evaluation of genetic parameters. *Euphytica*, 74 (1-2): 121-128.

Montagnon C., **Leroy T.**, Cilas C., Legnaté H., Charrier A (2008). Heterozygous genotypes are efficient testers for assessing between-population combining ability in the reciprocal recurrent selection of *Coffea canephora*. *Euphytica*, 160: 101-110.

Montagnon C, **Leroy T**, Cilas C, Charrier A (2003). Heritability of *Coffea canephora* yield estimated from several mating designs. *Euphytica*, 133: 209-218.

Montagnon C, Guyot B, Cilas C, **Leroy T** (1998a). Genetic parameters of several biochemical compounds from green coffee, *Coffea canephora*. *Plant Breeding*, 117: 576-578.

Montagnon C, **Leroy T**, Eskes AB (1998b). Amélioration variétale de *Coffea canephora*. I. Critères et méthodes de sélection. Plantations, recherche, développement, 5(1): 18-33.

Montagnon C, **Leroy T**, Eskes AB (1998c). Amélioration variétale de *Coffea canephora*. II. Les programmes de sélection et leurs résultats. Plantations, recherche, développement, 5(2): 18-31.

## ***C Amélioration des caféiers par transfert de gènes***

Après la présentation de ma thèse en juin 1993, une opportunité s'est présentée pour travailler sur la transformation génétique des caféiers en collaboration avec Nestlé, et plus particulièrement le centre de Recherches de Tours. Dans la continuité des travaux d'amélioration classique, l'utilisation du transfert de gènes pour l'amélioration ponctuelle d'un caractère ciblé est une voie intéressante.

Cette étape dans mes travaux a nécessité une formation initiale théorique et pratique pour acquérir les rudiments indispensables dans les domaines de la culture in vitro et de la biologie moléculaire. Le volet théorique a été acquis en suivant le DEA « Bases de la production végétale » sous la Direction du Professeur Macheix (octobre à décembre 1993). La partie pratique été réalisée dans les laboratoires du Cirad, tant en culture in vitro qu'en techniques de transformation génétique (Marc Berthouly et Thierry Legavre).

La collaboration entre le Cirad et Nestlé avait pour but de mettre au point un système reproductible de transformation génétique du caféier, et de mettre au champ des plantes transgéniques pour analyser l'ensemble du processus en conditions réelles. Une équipe de travail mixte Cirad-Nestlé dont j'assumais la coordination a été constituée pour mener à bien ces travaux :

- réalisation des constructions génétiques (M. Royer en stage post doctoral, Cirad),
- test de protéines insecticides de *Bacillus thuringiensis* (équipe de R. Frutos, Cirad),
- bio essais de protéines insecticides de *B. thuringiensis* sur les parasites du caféier (C. Fenouillet et I. Jourdan, Cirad),
- appui en culture in vitro (M. Berthouly, Cirad ; J-P. Ducos, Nestlé),
- réalisation des travaux de transformation (J. Spiral et S. Lechat, Nestlé),
- analyse moléculaire des plantes (M. Paillard, P. Marraccini et AM Henry, Nestlé),
- bio essais des plantes transformées (équipe de R. Philippe, Cirad),
- réalisation du dossier de dissémination des plantes pour la CGB (A-M. Henry en stage post doctoral, Nestlé),
- mise au champ et suivi des plantes (équipe de J-L. Pradon en Guyane).

Détaché sur le centre de recherches Nestlé à Tours d'avril 1994 à juin 2000, j'ai réalisé directement toutes les expérimentations de transformation dans les laboratoires de Nestlé, l'analyse moléculaire des plantes puis la rédaction du dossier pour la mise au champ des plantes, avant d'en assurer le suivi en Guyane jusqu'à la fin de l'essai en 2004.



## 1 Quelles constructions génétiques ?

Même si les travaux revêtaient un caractère prospectif pour la transformation génétique du caféier, le but de mise au champ des plantes nous a amené à considérer la réalisation de constructions génétiques d'intérêt pour l'amélioration des caféiers.

Une thèse d'un collègue brésilien (Oliveiro Gueireiro) se terminait. Il avait travaillé sur la mineuse des feuilles du caféier *Leucoptera coffeella*, parasite d'importance sur le caféier au Brésil et en Afrique de l'Est, et avait mis en évidence par des tests in vitro et sur feuilles que le gène de *Bacillus thuringiensis cry1Ac* était efficace contre ce lépidoptère parasite (Gueireiro et al. 1998). D'autre part, des travaux d'une autre thèse sur la transformation génétique du caféier par électroporation ou biolistique avaient permis d'identifier des agents sélectifs utilisables pour sélectionner les cellules transformées de caféier. Les herbicides bialaphos et chlorsulfuron et l'antibiotique hygromycine semblaient utilisables (Van Boxtel et al. 1997). Enfin, un promoteur d'élongation d'*Arabidopsis thaliana* avait été mis en évidence comme très actif sur caféier par ce même chercheur. C'est à partir de cette base de travail, et des travaux préliminaires de Nestlé sur la transformation du caféier par *Agrobacterium rhizogenes* (Spiral et al. 1993) que nos travaux ont débuté.

Dans notre démarche vers la production de plantes transgéniques au champ, les différents gènes et méthodes possibles pour les constructions génétiques ont été passés en revue :

- la méthode de transformation par *Agrobacterium tumefaciens* a été préférée en raison de la « propreté » des insertions et du faible nombre de copies inséré,
- les gènes de résistance aux herbicides (bialaphos et chlorsulfuron) ont été préférés au gène de résistance à l'hygromycine,
- un promoteur d'élongation d'*Arabidopsis thaliana* a été utilisé pour le gène d'intérêt de résistance à la mineuse des feuilles.

A partir de ces données, six constructions ont été faites, chacune composée de trois gènes :

- le gène *uidA* codant pour la coloration bleue des plantes transformées (test GUS),
- le gène de résistance à l'herbicide (chlorsulfuron ou bialaphos),
- le gène de résistance à la mineuse des feuilles *cry1Ac* sous sa forme native, ou synthétique pour une meilleure expression dans les dicotylédones (Sardana et al. 1996).

Les six constructions génétiques ont ainsi été introduites dans 4 bactéries (trois de *A. tumefaciens* et une de *A. rhizogenes*) et testées en transformation génétique.

A côté de ces travaux sur la mineuse des feuilles, parasite d'intérêt limité à quelques pays et conditions de culture, l'équipe et ses partenaires Cirad ont développé des recherches de souches de *B. thuringiensis* (*B.t.*) pour d'autres parasites d'importance pour le caféier, en particulier le scolyte des baies. Cet insecte, présent dans la plupart des pays producteurs sur les deux espèces de caféier cultivé, est un coléoptère qui attaque les fruits, et fait des dégâts très importants pouvant aller jusqu'à 50% de chute de production. En collaboration avec l'équipe de Roger Frutos, spécialisée dans la recherche de souches *B.t.* actives, un grand nombre de souches potentiellement actives contre les coléoptères a été testé sur des larves de scolyte issues d'un élevage au Cirad. Quelques souches ont donné des résultats intéressants, mais il a ensuite été impossible d'isoler les gènes responsables de cette activité.

En partenariat avec des collègues indiens du CCRI (Central Coffee Research Institute), des bio essais ont été faits sur le scolyte des baies mais aussi sur le White Stem Borer (WSB), parasite spécifique des troncs du caféier qui fait des ravages importants en Inde. Sur ce dernier parasite, deux souches de *B.t.* se sont montrées actives, et ont permis d'isoler deux gènes *cry35* et *cry36* actifs. Des plants transgéniques avec ces gènes n'ont pas encore été produits.

## 2 Obtention des plantes transformées

Nous avons ensuite procédé par étapes à partir des données disponibles sur la transformation génétique du café chez Nestlé et des travaux récents du Cirad, en utilisant des génotypes des deux espèces cultivées (*C. canephora* et *C. arabica*) :

- explant de travail : embryons somatiques ou cal embryogène en milieu liquide ou solide, obtention de matériel à transformer pour des génotypes divers,
- doses d'herbicide pour la sélection des cellules transformées,
- conditions de co-culture bactéries/explant, conditions de sélection des transformants,
- isolement des transformants et régénération d'embryons et de plantes transgéniques.

Les mises au point, effectuées dans le laboratoire de Nestlé, se sont poursuivies pendant deux ans. De nombreux contacts ont été pris avec des personnes travaillant sur ces thématiques (INRA de Versailles, INRA d'Orléans). Les premiers transformants ont été obtenus en 1996 par transformation d'un génotype de *C. canephora*, le clone 126, par *A. rhizogenes* avec le gène de résistance au bialaphos et le gène *cry* natif. Ces premiers transformants nous ont permis de préciser certains points du protocole, et dès 1997, un grand nombre de transformants a été obtenu par transformation avec *A. tumefaciens*, avec le gène synthétique de *cryIAc* et le gène de résistance au chlorsulfuron. Dès 1998, un nombre important de plantules transgéniques, positives au test coloré GUS étaient disponibles pour plusieurs génotypes de *C. arabica* et *C. canephora*. La transformation de cal embryogène avec la souche de *A. tumefaciens* LBA4404 s'est avérée la plus performante dans nos expériences.

## 3 Evaluation des plantes

Après cette étape d'obtention de cal transgénique puis de plantules, l'étape suivante a consisté à analyser les plantes obtenues. Cette analyse s'est effectuée en plusieurs étapes :

- vérification par test coloré gus de la nature transformée et non chimérique des plantes (plusieurs feuilles testées),
- analyse moléculaire par PCR et Southern blotting,
- analyse protéique par Western blotting,
- évaluation de la résistance effective à l'insecte par des bio essais sur plantes en laboratoire.

Dans la démarche vers la mise au champ de caféiers transgéniques, il était important de faire des choix dans les plantes à analyser au champ. Nous avons choisi la Guyane française, où le Cirad a une station de recherches en propre, et où les règles de la CGB (Commission du Génie

Biomoléculaire) française s'appliquaient. Les caféiers de l'espèce *C. canephora* étant les seuls à se développer normalement en climat guyanais, nous avons choisi de mettre au champ une soixantaine de transformants issus d'un même génotype de *C. canephora*. Ainsi, il serait possible d'évaluer avec précision l'effet de la transformation génétique pour un génotype donné.

Les analyses moléculaires ont donc été lancées sur près de 80 événements de transformation du clone de *C. canephora* 126 obtenus avec différentes constructions (gène synthétique ou natif), différents gènes de sélection (bialaphos ou chlorsulfuron) mais également après transformation avec *A. rhizogenes* ou *A. tumefaciens*. J'ai effectué les analyses moléculaires proprement dites (PCR, Southern and Western blotting) en m'appuyant sur les compétences du laboratoire de Nestlé et avec l'appui d'un étudiant en post doctorat (A. M. Henry). J'ai coordonné également les analyses de la tolérance des plantes aux insectes faites à Montpellier par R. Philippe et son équipe. Lors de cette phase d'analyse des plantes et de montage du dossier pour la mise au champ, j'ai passé près de 25% de mon temps à Montpellier pour coordonner les travaux de l'équipe dont j'étais responsable, qui travaillait aussi sur la recherche de protéines insecticides et le piégeage du scolyte des baies des caféiers.

Le bilan des analyses moléculaires et par bio essais des plantes a permis de faire le choix des plantes à mettre au champ, un dossier a été monté pour la CGB en 1998 (Leroy, 1998) pour une mise au champ prévue en 1999. Le Tableau 4 présente quelques résultats des analyses par Southern blotting d'une partie des plantes transformées mises au champ. Dans notre stratégie de connaissance des caféiers transgéniques, nous avons choisi de mettre aussi au champ quelques caféiers transformés avec *A. rhizogenes* présentant ou non le phénotype Hairy root, et des plantes transformées avec le gène synthétique ou natif, ayant ou non intégré le transgène. En effet, dans une optique générale d'amélioration des caféiers pour les caractères de résistance ou de qualité, il nous semblait intéressant de pouvoir mener une étude aussi complète que possible.

L'obtention et l'analyse des plantes transformées ont fait l'objet de nombreuses communications (Leroy et al. 1998) et d'une publication dans Plant Cell Reports (Leroy et al. 2000). La figure 4 présente des cals et des plantes transgéniques obtenues par notre équipe, ainsi qu'une illustration des bio essais réalisés à Montpellier avec des mineuses des feuilles sur des plantes transgéniques.

**Tableau 4 - Résultats des hybridations Southern avec la sonde promoteur ou la sonde GUS pour quelques transformants.** Le nombre de copies a été estimé grâce à la comparaison de l'intensité des bandes des témoins plasmidiques et des fragments internes de l'ADN-T intégré, mais aussi par le nombre de bandes correspondant aux fragments chevauchants la bordure droite. + : résultats conformes aux valeurs attendues pour les constructions génétiques utilisées. bd : intégration jusqu'au site BglII au delà de la bordure droite. nd : non déterminé. ni : résultat non interprétable.

Nom des plantes	N°attribué	Digestion 1	Digestion 2	Digestion 3	Nombre de copies	Remarques
<b>Plantes transformées par A. rhizogenes</b>						
MT 25 B2 (B2/A4)	43				1	Sonde GUS
MT 8 BTr (B2/A4)	1	+	+	+	2	Sonde Gus
MT 25 B1 (B2/A4)	4				2	Sonde Gus
MT 8 B1(C2/A4)	41				1	
MT 8 B9 (C2/A4)	2	Nd	+	nd	1 à 5	Sonde GUS
MT 27 B1 (C3/A4)	5	+	+	+	1	
MT 27 B2 (C3/A4)	60	+	+	+	2	1 copie avec la sonde GUS
MT 27 B3 (C3/A4)	61	+	3501	nd	1	bd
<b>Plantes transformées par A. tumefaciens</b>						
MT28 36 (C3/LBA)	11	+	+	+	1	
MT28 18	76				5	expliqué uniquement par l'intensité des bandes
MT28 9	8	+	+	+	1	
MT28 11	10	+	3501	nd	1	bd
MT28 A2	12	2 copies +	1 copie +	1 copie +	1 ou 2	Sonde GUS
MT28 30	22	+	+	+	2	
MT28 32	23	+	+	+	3	
MT28 A13	26	+	+	+	1	Sonde GUS
MT 33 17B	30	+	+	+	1	2 copies sonde GUS
MT 36 19	34	+	+	+	1	
MT28 2	45	+	+	+	1	
MT28 35	48	+	+	+	1	
MT28 A4	49	+	nd	+	1	
MT28 A8	50	+	+	+	1	
MT 33 16	51	+	+	+	1	Sonde GUS
MT 36 26	56	+	+	+	1	Sonde GUS
MT 20 B1	59	+	3501	nd	1	bd
MT28 5	62	+	+	+	3 ou 4	
MT28 6	63	+	+	+	3 ou 4	
MT28 8	64	+	+	+	1	
MT28 13	65	+	+	nd	1	Sonde GUS
MT28 A28	66	+	+	+	1	
MT28 B21	67	+	2x3501	+	1	bd
MT33 22	68	+	+	+	1	
MT33 23	69	Ni	+	+	4	

## 4 L'essai au champ

L'essai au champ a été installé avec un an de retard en 2000, et sa fin brutale en 2004 (arrachage des plantes par des inconnus) n'a pas permis d'en tirer toutes les informations possibles. Néanmoins, ces quatre années d'observation ont permis un certain nombre d'enseignements intéressants pour l'amélioration des caféiers. A la demande de la CGB, un dispositif très important de parcelles pièges à pollen avait été mis en place pour piéger les pollens transgéniques et connaître avec précision les distances de dissémination de celui-ci. Un certain nombre de dispositifs originaux ont été mis en place lors de cette expérimentation : layon dans la forêt sous les vents dominants avec des bosquets pièges à pollen sur 1 km, suivi de populations d'abeilles, lâcher de mineuses des feuilles dans la parcelle à intervalles réguliers.

Nous avons mis en évidence la conformité agronomique des plantes transformées par *A. tumefaciens* ou par *A. rhizogenes* qui ne présentaient pas le phénotype Hairy root (les autres sont mortes au cours de la première année). Aucune influence des plantes transgéniques sur les insectes non cibles présents dans les caféières n'a été observé. Ces études ont nécessité la mise en place de ruches dans les parcelles, et le suivi des populations d'abeilles, ainsi qu'un suivi de la composition des miels. Afin d'évaluer la résistance effective des plantes à l'insecte cible, et compte tenu de la faible pression du parasite dans la parcelle d'expérimentation, nous avons procédé à des lâchers de mineuse des feuilles dans la parcelle à partir d'un élevage mis en place en Guyane avec la souche locale de mineuses. Les résultats des comptages de mines sur les plantes montre qu'une part importante des événements de transformation mis au champ présente une bonne résistance au parasite dans les conditions de la Guyane (Perthuis et al 2005). Les analyses de dispersion du pollen n'ont pu être menées à bien en raison de la fin prématurée de l'essai.

## 5 Partenariats et perspectives

Ces travaux sur la transformation génétique des caféiers, conduits jusqu'à l'essai au champ, ont constitué une première sur le caféier à travers le monde. Cela nous a permis de nous positionner au niveau mondial sur ces problématiques pour les caféiers. Les résultats obtenus ont été présentés dans de nombreux congrès sur le café, les gènes *B.t.*, la pathologie des invertébrés ou les biotechnologies. Nous avons rédigé trois chapitres de livres sur la transformation génétique (Spiral et al. 1999 ; Leroy et Dufour 2004 ; Leroy et al. 2006) et l'équipe a participé à des réunions, rapports ou enseignements sur le sujet.

D'autre part, ces travaux novateurs nous ont permis de nouer des collaborations internationales. D'une part, nous avons pu monter de projets avec des collègues indiens sur des parasites spécifiques et partager avec eux notre expérience de la transformation génétique pour la création de plantes transgéniques et leur gestion au champ (Surekha et al. 2002). D'autre part, les premiers contacts avec le Brésil se sont noués sur les problématiques de la transformation génétique en 1998. Par la suite, les collaborations se sont élargies à la physiologie moléculaire puis à la génétique moléculaire, comme nous le verrons.



Figure 4 - Transformation génétique des caféiers.

A : cal transgénique(1) et embryons somatiques (2) poussant à partir d'un embryon nécrosé (3) sur milieu sélectif.

B : cal transgénique testé par coloration GUS.

C : plantes transformées de *C. canephora* (0 à 12 mois après transfert en serre de gauche à droite).

D : plantes transformées de *C. arabica* (et témoin non transformé à droite) 6 mois après le transfert en serre.

E : plante témoin non transformée, sensible à la mineuse des feuilles

F : plantes transformées avec le gène *cryIAc*, exposées à l'insecte au même moment. La plante 1 est sensible, la plante 2 est résistante aux attaques de l'insecte.

Ces compétences sur la transformation génétique et la mise au champ de ces plantes ont donc constitué une étape importante pour élargir mes compétences en sélection classique vers l'utilisation des outils de la biotechnologie pour améliorer les caféiers, et ont ouvert à l'équipe de nouvelles voies de collaboration en Asie et en Amérique.

## Principales publications de l'équipe

**Leroy, T** (1998). Dossier de demande d'expérimentation au champ de caféiers génétiquement modifiés auprès de la Commission du Génie Biomoléculaire.

**Leroy T**, Alpizar E, Dufour M, Etienne H (2006) Coffee (*Coffea* sp.): In *Agrobacterium protocols*, Second edition: volume II (Methods in molecular biology, vol.44), K Wang (ed.), pp 191-208. Humana press, Totowa, NJ, USA.

**Leroy T**; Dufour M (2004). *Coffea spp.* Genetic transformation. In *Transgenic crops of the world- Essential protocols*. IS Curtis (ed.), pp 159-170. Kluwer Academic Publishers, London, Great Britain.

**Leroy T**, Henry AM, Royer M, Altosaar I, Frutos R, Duris D, Philippe R (2000). Genetically modified coffee plants expressing the *Bacillus thuringiensis* cry1Ac gene for resistance to leaf miner. *Plant Cell Reports*, 19 (4): 382-389.

**Leroy T**, Paillard M, Royer M, Spiral J, Berthouly M, Tessereau S, Legavre T, Altoosaar I (1998). Introduction de gènes d'intérêt agronomique dans l'espèce *Coffea canephora* Pierre par transformation avec *Agrobacterium* sp. In 17<sup>th</sup> International Conference on Coffee Science (ASIC), Nairobi, Kenya, July 1997, 20-25, pp. 439-445. ASIC, Paris, France.

Perthuis B, Pradon JL, Montagnon C, Dufour M, **Leroy T** (2005). Stable resistance against leaf miner *Leucoptera coffeella* expressed by genetically transformed *Coffea canephora* in a pluriannual field experiment in French Guiana. *Euphytica*, 144: 321-329.

Spiral J, **Leroy T**, Paillard M, Petiard V (1999). Transgenic Coffee (*Coffea* species). In *Biotechnology in Agriculture and Forestry*, Vol. 44 *Transgenic trees*, YPS Bajaj (ed.), pp 55-76. Springer-Verlag Berlin Heidelberg, Germany.

Surekha K, Royer M, Naidu R, Vassal JM, Philippe R, Jourdan I, Fenouillet C, **Leroy T**, Dufour M (2002). Bioassay of *Bacillus thuringiensis* toxins against two major coffee pests, i.e. coffee berry borer (*Hypothenemus hampei*) and coffee white stem borer (*Xylotrechus quadripes*). In SIP, Annual Meeting of the Society for Invertebrate Pathology, Foz do Iguassu, Brazil, August 2002, 18-23, p.85. SIP, Irvine, USA.

## ***D Génomique de la qualité des cafés***

L'utilisation des outils de la génomique pour l'amélioration des caféiers et des cafés dans notre équipe a découlé naturellement des travaux précédents en génétique classique et transformation génétique. Si cette dernière technique permet d'introduire spécifiquement tel ou tel gène dans les caféiers, il convient de déterminer les gènes et allèles d'intérêt pour les caractères de productivité, de qualité ou de tolérance aux stress. D'autre part, afin d'améliorer l'efficacité de la sélection classique, la mise en évidence de marqueurs liés aux gènes d'intérêt utilisables en sélection précoce est un outil puissant.

Nous avons donc initié un certain nombre de projets en partenariat avec nos collègues africains et brésiliens sur l'amélioration de la qualité des cafés en utilisant les outils de la génomique. Pour mener à bien ces travaux, l'équipe s'est renforcée avec l'arrivée en 2001

d'un physiologiste moléculaire (P. Marraccini) puis en 2004 d'un généticien moléculaire (David Pot) tous les deux affectés au Brésil. Mon travail de coordination a donc pris une dimension nouvelle avec la présence de collègues de l'équipe au Brésil, les travaux menés en Ouganda sur la diversité et l'intégration dans l'équipe des chercheurs de Guyane travaillant sur le café.

Les travaux reposaient en effet sur quatre terrains avec lesquels nous avons des contacts privilégiés pour l'accès au matériel végétal et qui étaient impliqués dans des projets communs:

- la Côte d'Ivoire, lieu d'implantation de tous les essais de sélection récurrente, qui possède la collection la plus importante en termes d'espèces sauvages et de *C. canephora*,
- l'Ouganda, partenaire dans trois projets relatifs à la qualité des cafés, et pays intéressant pour la diversité génétique de ses caféiers,
- le Brésil et nos partenaires du IAPAR (Londrina, Parana), du CENARGEN (Brasilia, DF) et de l'IAC (Campinas, Sao Paulo) qui possèdent des collections importantes de *C. arabica* et ont des équipes très importantes qui travaillent sur les caféiers,
- la Guyane, où le Cirad dispose d'une collection de toutes les espèces de caféier.

Après une approche globale de la productivité et de la qualité du produit, avec l'analyse de la transmission des caractères et de la diversité pour les paramètres de la qualité, la démarche suivie était de décomposer les mécanismes moléculaires qui conduisent à la formation des composés impliqués dans la qualité des cafés (sucres, lipides, ...).

Avant que ne débute la collaboration avec les équipes brésiliennes, des contacts avaient été pris au niveau méthodologique avec l'INP- ENSA Toulouse (JC Pech) sur la maturation des fruits en général. Les travaux déjà entrepris sur le melon ont servi de base à nos travaux sur le caféier pour les études de maturation. Une ATP (Activité Thématique Programmée) Cirad avait été présentée fin 2001 avec des recherches sur le caféier, le bananier et l'ananas pour la maturation des fruits en relation avec la qualité des produits. Elle n'a pas été retenue mais la réflexion a été à la base des travaux entrepris par la suite.

## **1 Mise en place des outils moléculaires**

Dans le cadre d'un projet avec le géoscope, des banques enrichies en microsatellites ont été faites pour plusieurs plantes. Une banque de marqueurs microsatellites café a été créée, près de 800 séquences microsatellites ont été identifiées (Dufour et al. 2001). Ces marqueurs ont été utilisés dans les travaux de cartographie génétique et de diversité pour l'espèce *C. canephora* et le genre *Coffea* (Poncet et al. 2007 ; Cubry et al. 2008).

J'ai construit la banque BAC de *C. canephora* en 2002, afin de disposer de cet outil pour les travaux de physiologie moléculaire débutés au Brésil. Le choix du génotype à utiliser pour la construction de la banque s'est porté sur le clone 126, reconnu comme produisant un café de bonne qualité, pour augmenter les chances de repérer facilement les allèles favorables des gènes codant pour les caractères de qualité. La construction de cette banque a permis de disposer de 55000 clones BAC avec une longueur moyenne des inserts de 135 Kb (Leroy et al 2005). Les collègues de l'IRD ont réalisé en même temps que nous une banque BAC sur *C. arabica* (Noir et al, 2005). Nous avons donc utilisé les mêmes marqueurs d'une carte



génétique récente (Lashermes et al. 2001) et positionnés sur 10 des 11 groupes de liaison (chromosomes) de *C. canephora* pour valider notre banque. 13 marqueurs utilisés étaient simple copie sur arabica, et 4 étaient dupliqués. Nous avons ainsi confirmé que notre banque représentait en théorie l'équivalent de 9 copies du génome haploïde. Cette banque, outre son utilisation dans les travaux présentés par la suite, a aussi positionné l'équipe en génomique structurale des caféiers. Un collègue brésilien est venu quelques semaines en stage dans notre équipe pour préparer la construction de sa propre banque BAC sur un hybride interspécifique entre *C. arabica* et *C. canephora*. Le séquençage de BACs ciblé a également permis de mettre en évidence un nombre important de marqueurs microsatellites cartographiables sur notre carte génétique, et positionnés à proximité de zones d'intérêt pour la qualité du café. Cette banque génomique a également été utilisée pour des travaux avec des collègues de l'IRD (Bustamante et al. 2007) et de Nestlé.

Nous avons fait le choix de ne pas développer nous même de ressources EST dans l'équipe. Ce choix a été guidé par le fait que des ressources importantes étaient disponibles ou en voie de l'être. Les équipes de Nestlé/Université de Cornell d'une part et les équipes brésiliennes d'autre part avaient créé des ressources importantes (respectivement 45.000 et 150.000 clones EST) auxquelles nous avons eu accès dans le cadre de nos collaborations avec ces laboratoires.

## **2 Mécanismes moléculaires de la construction de la qualité**

Les travaux ont débuté dans le cadre d'une collaboration entre un laboratoire brésilien (le IAPAR) et le Cirad en 2001. La collaboration a pour titre « Amélioration de la qualité des cafés ». C'est un projet d'amélioration de la qualité du café (Arabica) avec deux axes de travail :

- Le déterminisme des facteurs physiologiques et moléculaires pour la maturation des fruits et la qualité,
- L'utilisation de la transformation génétique pour le caféier.

Outre le détachement d'un chercheur du Cirad (P. Marraccini) au IAPAR, le projet a permis la venue au Cirad d'un chercheur brésilien pour travailler sur l'évaluation de la qualité par dégustation et par absorption dans l'Infra rouge (NIRS). Ce projet avait pour but :

- d'acquérir une bonne connaissance des déterminismes de la qualité des cafés par physiologie moléculaire,
- de créer un pôle de compétences sur la qualité des cafés et les analyses sensorielles avec les laboratoires du IAPAR et du Cirad.

Les travaux ont débuté sur le métabolisme des sucres. Ce choix a été guidé par les travaux déjà effectués ou en cours dans les autres laboratoires : acides chlorogéniques et caféine dans les laboratoires de l'IRD, sucres complexes dans les laboratoires de Nestlé. D'autre part, la concentration en saccharose des baies de *C. arabica* est plus importante (5 à 9%) que celle de *C. canephora* (4 à 7%). Or, les sucres simples et en particulier le saccharose sont des précurseurs importants des arômes dans les cafés. En effet, le saccharose se dégrade rapidement durant la torréfaction, donnant des molécules qui interagissent avec les acides aminés (réaction de Maillard) pour donner des composés comme les furanes qui sont considérés comme des contributeurs importants dans la flaveur des cafés.

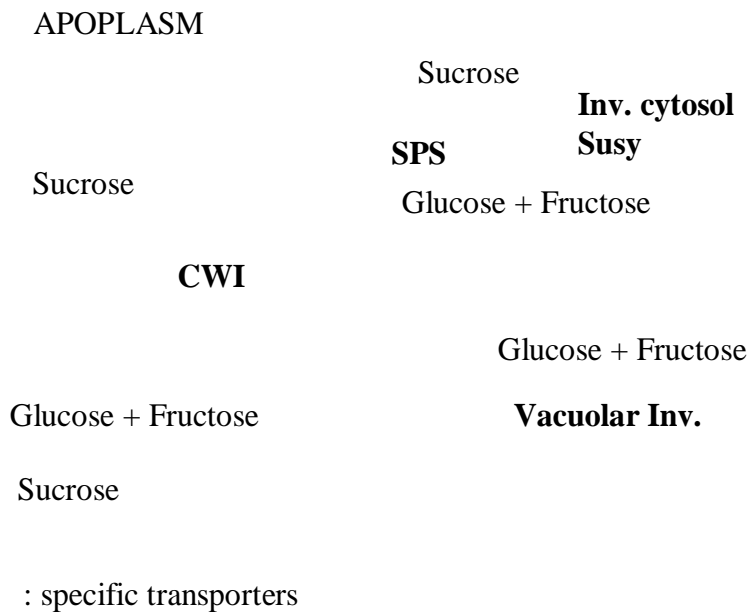


Figure 5 - Métabolisme des sucres dans la cellule. Les enzymes impliqués sont les invertases vacuolaires (Vacuolar Inv, INV) et pariétales (Cell Wall Invertases, CWI), les saccharoses synthétases (Susy) et la saccharose phosphate synthétase (SPS).

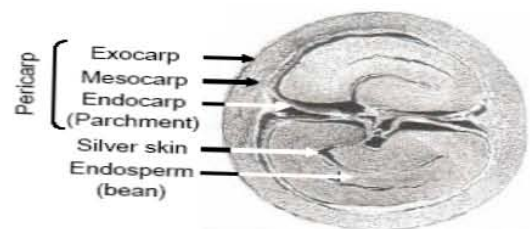


Figure 6 - Tissus du fruit de caféier.

Les principales enzymes impliquées dans le métabolisme des sucres (invertases, saccharose synthétases et saccharose phosphate synthétases) sont connues mais leur fonctionnement lors de la maturation de fruits de café et dans les différentes parties du fruit (péricarpe, péricarpe, endosperme) n'était pas connu au début de notre projet. Les travaux ont été faits sur un génotype de *C. arabica* cultivé au IAPAR, et ont permis la présentation d'une thèse (Clara Geromel) à l'UNICAMP (Université de Campinas, Sao Paulo), partenaire du Cirad et du IAPAR pour la partie biochimique de ce travail.

Le métabolisme des sucres, en général, et les mécanismes généraux dans la cellule pour le métabolisme des sucres sont présentés dans la Figure 5. Le rôle des différentes enzymes y est présenté globalement. Le développement de la baie de caféier dure 250 jours après la floraison, les différents tissus des fruits sont présentés sur la Figure 6. La fève de café est constituée de l'endosperme dont l'expansion au cours de la maturation se fait aux dépens du péricarpe réduit dans la baie mûre à une pellicule argentée. Nos travaux ont permis d'analyser :

- l'évolution du saccharose et des sucres réducteurs au cours de la maturation,
- l'évolution des différentes enzymes au cours de la maturation pour les trois tissus principaux.

Les saccharose synthétases sont une famille multigénique importante (Baud et al, 2004). Deux isoformes de saccharose synthétases ont été mises en évidence dans nos études. L'une (*Ca susy1* pour *Coffea arabica susy1*) s'exprime surtout dans les premiers stades de développement et la croissance du péricarpe, alors que l'autre (*Casusy2*) est impliquée dans les stades finaux du développement et le remplissage du grain (endosperme). Cette dernière isoforme pourrait donc être un gène majeur responsable de l'accumulation du saccharose dans la graine de café (Geromel et al. 2006). Des études complémentaires chez des caféiers en situation d'ombrage et de plein soleil ont montré par ailleurs le rôle important de la saccharose phosphate synthétase et les modifications qui peuvent intervenir pour les saccharose synthétases en situation d'ombrage (Geromel et al. 2008a). Des travaux effectués sur une autre espèce de caféier *C. racemosa* (Geromel et al. 2008b) donnent des résultats légèrement différents.

Des séquences partielles de cDNA des principales enzymes du métabolisme des sucres ont été utilisées lors de ces travaux pour évaluer l'expression des gènes lors de la maturation. Ces sondes ont été hybridées sur la banque BAC de *C. canephora*. Ainsi, il a pu être mis en évidence que les invertases vacuolaires étaient représentées par une seule copie du gène dans le génome. En revanche, le gène de *susy1* et l'invertase pariétale CWI sont présents en 2 ou 3 copies dans le génome de *C. canephora*, alors que *susy2* n'est présent qu'en une seule copie. La banque BAC a aussi permis, par séquençage des clones de la banque qui s'hybridaient avec la sonde *susy1*, d'avoir la séquence complète du gène et du promoteur du gène *Ccsusy1* (*Coffea canephora susy1*).

Les travaux entrepris sur les sucres ont donc permis de mettre en évidence l'importance des saccharose synthétases dans l'accumulation des sucres dans la fève de café. Le rôle des invertases apparaît mineur pour cette accumulation de saccharose dans les fruits. Le rôle exact de la saccharose phosphate synthétase est en cours d'étude. Les analyses effectuées sur des arbres sous ombrage montrent cependant la complexité des phénomènes, puisqu'une expression très forte du gène *Casusy2* en fin de maturation ne s'accompagne pas d'augmentation significative de la teneur en sucres des arbres sous ombrage, pourtant

reconnus comme meilleurs du point de vue organoleptique. En conclusion, si les sucres constituent bien un déterminant important pour la qualité des cafés, leur héritabilité très moyenne, les différences faibles observées entre caféiers dans différentes conditions de culture, et le rôle encore peu connu de certaines enzymes comme les saccharose phosphate synthétases montrent la nécessité de poursuivre les études, en particulier par comparaison entre *C. arabica* et *C. canephora*.

Bien évidemment, les sucres ne constituent pas les seuls composés qui interviennent dans la qualité des cafés. Les lipides, autre grande famille de composés, ont aussi un rôle important pour la fixation des arômes et entrent dans la composition des graines à hauteur de 10% (*C. canephora*) à 15% (*C. arabica*). Nous avons donc débuté des analyses sur une famille de composés lipidiques, les diterpènes, qui sont impliqués dans la qualité du café, mais aussi avec un impact sur la santé (anti oxydant et augmentation du taux de cholestérol sanguin). Des travaux ont débuté dans notre équipe sur ces composés en s'intéressant à trois gènes de la chaîne de biosynthèse de ces composés : copalyl diphosphate synthétase (*cps*), kauren oxydase (*ko*) et kauren synthétase (*ks*) qui sont impliqués dans la synthèse du Ent-Kaurene, produit intermédiaire dans la synthèse des trois diterpènes connus du café le Cafestol, le Kahweol et le 16-O- Méthyl Cafestol.

Plusieurs stagiaires au Brésil et en France ont été encadrés sur ce sujet, en particulier un stagiaire de Master 1 qui a travaillé dans le cadre des travaux sur la qualité des Robustas ougandais. Ce projet portait sur les qualités comparées des différentes origines des caféiers cultivés en Ouganda, l'analyse des diterpènes a permis de discriminer des origines, mais les discriminations sont différentes suivant les composés. Les analyses des métabolismes des sucres et des diterpènes se poursuivent actuellement avec nos collègues du Brésil pour identifier des gènes d'intérêt impliqués dans l'élaboration de la qualité des cafés.

Par ailleurs, un projet ANR génoplate coordonné par le Cirad (B. Bertrand) avec la participation de collègues de l'IRD et de Nestlé a pour but de connaître l'évolution globale du transcriptome au cours de la maturation de fruits de caféier par hybridation de puces à ADN. Ce projet est en cours, et des résultats sont attendus fin 2009. Notre équipe travaillera sur un génotype de *C. arabica* introgressé cultivé au Brésil, et analysera de façon comparative l'évolution du transcriptome d'un génotype de *C. canephora* planté dans les mêmes conditions.

### **3 Carte génétique de *C. canephora***

La cartographie génétique et la recherche des QTL liés à la qualité du café a été entreprise dans le cadre d'un projet européen INCO IRD/Cirad/RCI/Ouganda/Espagne sur qualité des robustas africains. Ce projet sur 5 ans a mobilisé l'équipe à Montpellier, et a été soutenu par l'encadrement de nombreux stagiaires de DEA ou DESS qui ont travaillé sur la construction de la carte génétique (M. Berline, K. Avia), et sur l'analyse des déséquilibres de liaison entre marqueurs dans des populations définies par les études de diversité (K. Avia, P. Cubry). Ainsi, les travaux de cartographie proprement dits sont reliés avec les analyses de diversité pour définir des populations d'études dans lesquelles les liaisons entre les marqueurs sont connues afin de définir les stratégies d'amélioration intra et inter populations. Ce projet, en complément des travaux menés dans le cadre de l'ATP « déséquilibre de liaison » a permis à l'équipe de construire des compétences dans ce domaine.

Notre population de cartographie est un pseudo backcross (Guinéen x Congolais) recroisé par Guinéen comportant 248 individus plantés au champ sur la station de Divo en Côte d'Ivoire. La cartographie génétique proprement dite a permis de disposer fin 2007 d'une carte de plus de 200 marqueurs microsatellites répartis sur les 11 groupes de liaison de *C. canephora* avec une longueur totale de la carte proche de la longueur totale estimée du génome (de l'ordre de 1500 cM). Outre les marqueurs issus de notre banque enrichie (Dufour et al. 2001) et des banques travaillées par les collègues d'autres institutions (Combes et al. 2000 ; Rovelli et al. 2000), les collègues du laboratoire Nestlé à Tours nous ont cédé 200 séquences microsatellites déjà cartographiées sur leur descendance de *C. canephora*. La carte génétique est présentée sur la Figure 7. Le point intéressant est la cartographie de gènes et de séquences proches de gènes du métabolisme des sucres et des lipides :

- trois gènes sont cartographiés directement sur la carte : le gène *susy2* (indel dans un intron), le gène *koCI-2* (indel dans la partie 3'UTR du gène), le gène *susy1* (deux microsatellites dans la séquence promotrice et deux microsatellites dans la séquence du gène (introns et exons),
- des microsatellites repérés dans des séquençages terminaux de clones de la banque BAC ayant donné une hybridation avec des sondes des gènes *susy2*, invertases vacuolaires et pariétales (deux localisations sur la carte), *susy1* (trois localisations, donc 3 copies du gène), et *sps*.

Notre carte génétique constitue donc un outil complémentaire de nos analyses de physiologie moléculaire pour positionner sur le génome un certain nombre de gènes intéressant la qualité des cafés. Une grande majorité des marqueurs fournis par Nestlé sont positionnés sur les mêmes groupes de liaison que sur leur carte. Un travail de comparaison des cartes est en cours actuellement. Un marquage dense existe déjà pour la zone correspondant à certains gènes du métabolisme des sucres (*susy1* et *susy2*), et constitue une base pour les études d'association.

Plus récemment, à travers un projet de cartographie sur une descendance d'hybrides Arabusta (*C. canephora* tétraploïde croisé par *C. arabica*) planté à l'IAC (Institut Agronomique de Campinas), les compétences de l'équipe en termes de cartographie de marqueurs microsatellites ont donné lieu à un stage pré doctoral de deux mois (L. Ferreira du IAPAR) pour cartographier des marqueurs microsatellites sur cette carte « Arabica ». Ces travaux, outre la densification de la carte brésilienne à base de marqueurs AFLP, permettront de positionner un certain nombre de marqueurs liés à des gènes d'intérêt sur une carte tétraploïde et de faire de la cartographie comparée avec la carte diploïde de *C. canephora*. De plus, il sera possible de comparer le positionnement des QTL obtenus sur les deux cartes.

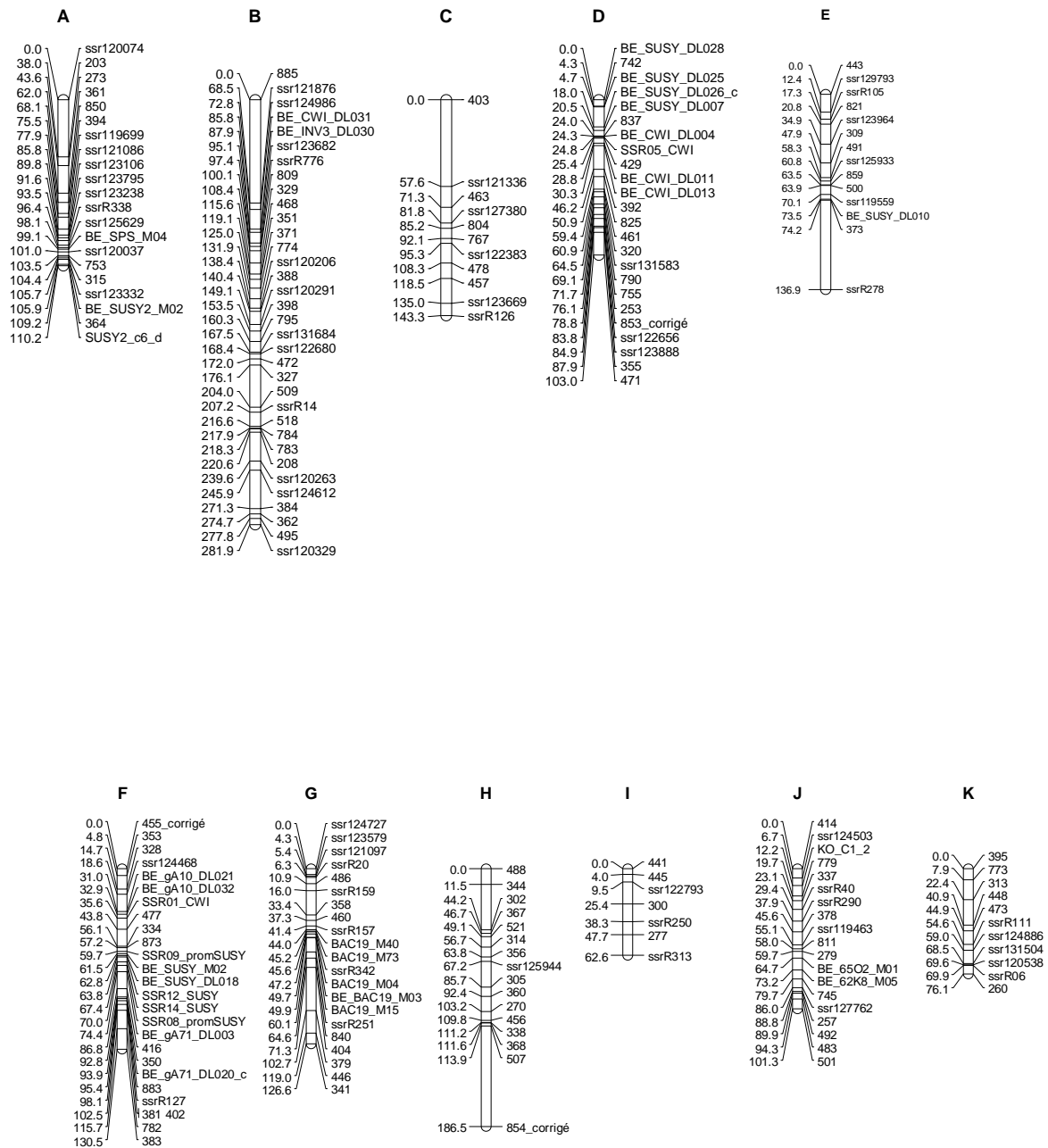


Figure 7 - Carte génétique de *C. canephora* : croisement de type pseudo backcross (Guinéen X Congolais) recroisé par Guinéen. 248 individus, 204 marqueurs cartographiés. Des marqueurs liés à des gènes (SUSY, SUSY2, CWI, INV, KO) ou à un clone BAC synthénique avec la tomate (BAC 19) sont indiqués.

L'existence de différentes cartes travaillées en commun (cartes Nestlé, Cirad et IAC) devrait permettre de positionner des QTL de manière assez précise. De plus, dans le cadre de l'ICGN (International Coffee Genome Network), j'assume la coordination du groupe « diversity and mapping ». Un état des lieux a été fait sur les cartes en cours pour les deux espèces cultivées, et un projet de carte « consensus » d'une descendance distribuée à tous les partenaires intéressés est en cours. Ce type de travail coordonné, difficile à mettre en place sur le caféier où les intérêts des pays producteurs (Colombie, Brésil, Inde,...) sont très forts, est très important pour faire progresser la mise en évidence de gènes d'intérêt chez les caféiers.

Ces premiers résultats encourageants de cartographie devront être développés pour cerner plus étroitement les zones du génome impliquées dans la variation de ces caractères. C'est pourquoi l'utilisation de la banque BAC pour connaître très précisément certaines zones, ou la réalisation d'un maillage de marqueurs plus denses en utilisant par exemple les AFLP ou les DART (projet déposé à la région Languedoc-Roussillon) devrait permettre de progresser dans cette voie. La cartographie de SNP (voir 4) pourrait permettre aussi une cartographie fine dans les zones d'intérêt.

L'analyse de QTL est en cours sur cette descendance et a permis de mettre en évidence des zones QTL pour certains caractères comme la taille des fèves ou l'acidité de la boisson (Tableau 5). Mais les zones QTL couvrent plusieurs dizaines de cM pour la plupart des caractères étudiés, les pourcentages d'explication restent faibles, ou ne sont valables que pour l'année considérée.

## **4 Diversité moléculaire**

Une recherche de SNP a été menée dans les séquences connues des gènes travaillés dans nos travaux ou mis en évidence dans les bases de données. Ces analyses ont constitué les travaux de DEA de S. Bouchet (métabolisme des sucres), et une partie des travaux de Master1 de N. Durand (diterpènes), ainsi qu'une partie des travaux en cours de D. Pot avec nos collègues du IAPAR. Cette recherche de SNP s'est faite par séquençage systématique de fragments dans les gènes des métabolismes concernés ou par recherche directe des séquences dans les banques d'EST. Les essais de génotypage automatique de SNP par d'autres méthodes comme l'ecotilling se sont soldés par des échecs en raison du trop grand nombre de polymorphismes dans nos échantillons.

Les SNPs ont été analysés pour l'ensemble du genre *Coffea* et pour les groupes identifiés de *C. canephora* pour une vingtaine de génotypes. Par séquençage manuel, 13000 pb ont été explorés pour 8 gènes codant pour 4 enzymes pour le métabolisme des sucres (Susy, Sps, CWI, INV). Pour le métabolisme des lipides, 3200 pb ont été explorées pour trois gènes (*ko*, *ks*, *cps*). Les résultats montrent un taux de polymorphismes de 1 SNP toutes les 82 pb pour les gènes du métabolisme des sucres et un SNP toutes les 35 pb pour le métabolisme des lipides. Les polymorphismes peuvent être situés sur les parties non codantes des gènes (3'UTR ou introns) ou dans les exons.

Tableau 5 – Identification de zones QTL par Interval Mapping dans une descendance de *C. canephora*. Pour chaque QTL, sont indiqués le groupe de liaison (A à K), le seuil de détection au niveau de chaque chromosome, et la zone correspondant à ce seuil avec les marqueurs qui la bordent, le LOD maximum observé et la variation phénotypique expliquée pour chaque caractère.

Traits	QTL name	Linkage Group	Confidence	Group (in cM)		Flanking markers		Maximum LOD	% explanation for the trait
				from	to	between	and		
2004_Aftertaste	2004_Aftertaste_2	D	1%	92.7	112.1	742	DL026	4.74	33
2005_Aftertaste	2005_Aftertaste_3	I	1%	0	11	ssrR313	277	4.95	19.3
2004_Pea	2004_Pea_8	F	1%	86.1	88.1	477	DL032	4.13	12.4
2004_Pea	2004_Pea_6	F	1%	43.8	69.3	416	334	6.3	17.7
2005_Pea	2005_Pea_6	J	1%	61.7	64.3	378	ssrR40	3.62	10.4
2005_Pea	2005_Pea_11	K	1%	61	81.7	313	773	4.59	17
2006_Pea	2006_Pea_7	J	1%	81.1	86.5	337	KO_C1_2	4.03	10.2
2006_Pea	2006_Pea_4	G	1%	18	33.2	341	404	3.35	11.7
2006_Pea	2006_Pea_2	F	1%	8	94.3	383	DL021	10.5	23
2005_lg(acid)	2005_lg(acid)_5	D	1%	128.2	131.2	DL025	DL028	7.97	54.7
2005_lg(acid)	2005_lg(acid)_2	A	1%	4	19	ssr120074	203	5.16	57
2005_lg(acid)	2005_lg(acid)_1	I	2%	58.5	59.5	445	441	4.68	31
2006_lg(acid)	2006_lg(acid)_2	I	1%	16.9	62.6	277	441	8.17	55.6
2004_size	2004_size_6	F	1%	104	124.4	DL021	353	5.55	13.3
2005_size	2005_size_2	A	1%	92.7	102.2	ssr123795	315	4.28	10.4
2005_size	2005_size_7	H	1%	38.2	89.8	305	344	5.88	14.6
2006_Bitterness	2006_Bitterness_2	I	1%	15.9	45.1	277	ssr122793	5.08	15.5
2005_Global	2005_Global_2	A	1%	70.3	75.7	850	394	4.02	16.5
2004_Flavour	2004_Flavour_2	D	1%	34.9	40.9	ssr122656	790	3.51	40.3
logcum03	logcum03_6	J	1%	88.5	96.6	779	414	3.58	10.2
logcum03	logcum03_2	D	1%	55.2	59.2	ssr131583	825	4.17	10.9
logcum03	logcum03_8	K	1%	37.9	44.9	448	313	3.53	15.3
logcum03	logcum03_4	I	1%	58	59.6	ssr122793	441	3.25	22.4
logcum06	logcum06_4	A	1%	76.7	76.7	394	ssr119699	3.62	13.2
logcum06	logcum06_6	K	1%	30.3	58	ssrR111	772	5.64	20
logcum04_06	logcum4_6_8	H	1%	39.9	81.7	448	773	3.34	10.3
logcum03_06	logcum3_6_3	A	1%	62.9	62.9	361	850	2.63	9.4
logcum03_06	logcum3_6_5	K	1%	44.9	55	448	772	3.21	9.6

Pea: percentage of pea berries; size: weight of 100 beans; lg(acid): logarithm transformation for acidity; logcumxx\_yy: logarithmic transformation of cumulated yield for years 20xx to 20yy



La part de polymorphismes dans les exons est faible (moins de 20% pour le métabolisme des sucres et de l'ordre de 50% pour le métabolisme des lipides). Pour les polymorphismes non synonymes (modification de la protéine) les situations sont très contrastées :

- le gène *susy1* ne présente pas du tout de polymorphisme non synonyme, mais des polymorphismes sont présents dans la séquence promotrice de ce gène,
- le gène *susy2*, responsable de l'accumulation du saccharose dans les fruits présente en revanche un taux de polymorphisme NS plus élevé que *susy1*,
- Il existe trois fois plus de polymorphismes non synonymes dans les gènes du métabolisme des diterpènes que dans ceux du métabolisme des sucres.

Ces résultats, associés à ceux obtenus en physiologie moléculaire, nous aident à définir une stratégie d'amélioration de la qualité des cafés. En effet le gène *susy1* est reconnu comme étant impliqué dans le métabolisme de base de la plante, est de ce fait soumis à de très fortes pressions de sélection. Il sera donc difficile de travailler directement sur ce gène pour l'amélioration. En revanche, il sera possible de travailler sur les séquences régulatrices de ce type de gène. Pour le gène *susy2*, il devrait être possible de mettre en évidence des variabilités dans nos populations qui pourraient directement être impliquées dans l'accumulation de saccharose dans les fruits. La cartographie de ces SNP et une analyse de QTL pourraient nous conduire à mettre en évidence des allèles favorables pour la sélection assistée par marqueurs. Quant aux gènes du métabolisme des diterpènes, moins soumis à des pressions sélectives que ceux du métabolisme des sucres, ils seront probablement plus faciles à travailler. Cependant les gènes travaillés sont en amont du métabolisme des diterpènes eux-mêmes, et doivent donc être considérés avec prudence.

Une étude des relations phylogénétiques entre les haplotypes déduits de nos analyses de polymorphisme a été effectuée chez l'espèce *C. canephora* et pour le genre *Coffea*. Contrairement à ce qui était attendu, la structuration mise en évidence par les marqueurs neutres et par les SNP pour les gènes soumis à forte sélection n'est pas différente. Pour le genre *Coffea*, l'arbre construit pour différents gènes (*susy*, *sps*) montre une structure du genre identique à celle observée avec les microsatellites (Figure 8 pour *susy2*). Ce qui est remarquable pour l'origine de l'espèce *C. arabica*, c'est le positionnement d'un haplotype avec le groupe de *C. canephora* et de l'autre avec le groupe d'Afrique centrale comprenant *C. eugenioides*. L'origine de *C. arabica* par croisement entre *C. canephora* et *C. eugenioides* est ainsi confortée. Pour la différenciation intra *C. canephora*, les résultats sont aussi identiques à ceux observés pour les marqueurs neutres.

Ces résultats nous confirment la faible domestication des caféiers et la structuration récente des populations chez *C. canephora*, conduisant à une structure de la diversité identique avec les deux types de marqueurs. Les résultats de ces études sur les polymorphismes des gènes du métabolisme des sucres et des lipides nous conduisent à utiliser ces polymorphismes pour les cartographier sur la carte génétique en cours de construction avec les marqueurs neutres, puis à essayer de les co-localiser avec des QTL de qualité. Ainsi, il pourrait être possible d'identifier les allèles favorables pour les caractères de qualité.

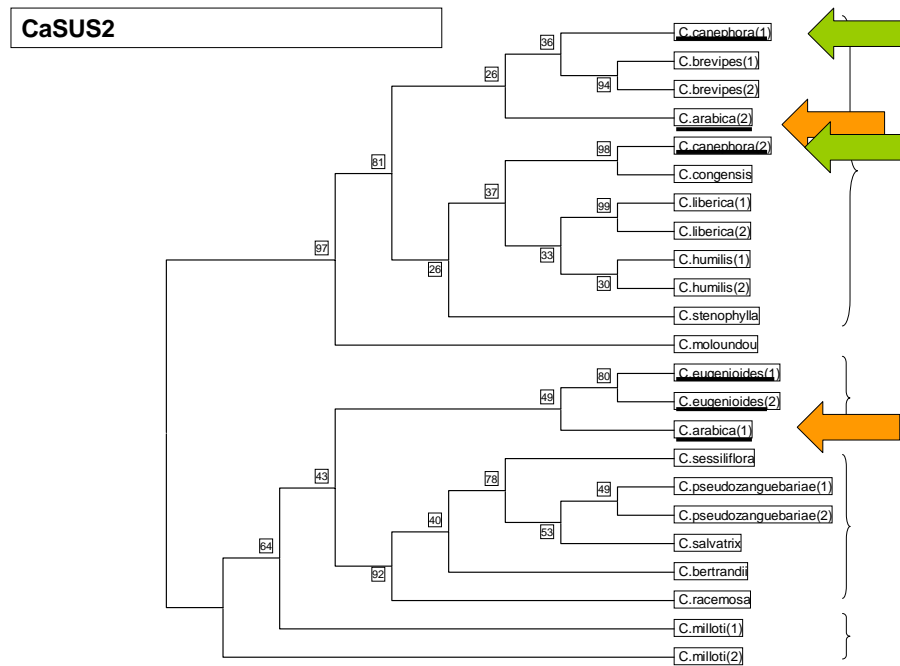


Figure 8 - Arbre phylogénétique construit à partir des polymorphismes du gène *susy2*. Un des haplotypes de *C. arabica* se groupe avec *C. canephora*, l'autre se groupe avec *C. eugenioides*, autre parent putatif de l'Arabica (Pot et al. 2006).

## 5 Analyses préliminaires du DL dans des populations de *C. canephora*

En complément des travaux de cartographie génétique, et dans la perspective de travaux en génétique d'association sur les populations de caféiers chez les partenaires, des analyses du déséquilibre de liaison entre les marqueurs dans nos populations d'intérêt sont menées. Ces analyses ont deux intérêts principaux :

- connaître la densité de marqueurs nécessaires sur le génome de *C. canephora* pour mener à bien les études d'association dans les populations analysées,
- éventuellement, à partir des connaissances sur la liaison moyenne entre marqueurs, permettre l'identification de marqueurs soumis à des forces de sélection particulières.

Ces travaux ont été développés dans le cadre des DEA de K. Avia et P. Cubry. Des résultats préliminaires montrent des intensités de liaison à grande distance pour certaines populations comme la population sauvage guinéenne Pélési. Le caractère isolé de cette population en milieu difficile pourrait expliquer en partie ces résultats. D'autre part, les populations cultivées comme les génotypes du groupe SG2 semblent présenter des liaisons entre marqueurs faibles même à courte distance. Pour ces travaux, seuls les marqueurs microsatellites sont analysés pour l'instant. Cependant, pour les gènes d'intérêt du métabolisme des sucres et des lipides, la cartographie de marqueurs SNP serait souhaitable pour densifier la carte dans des zones d'intérêt. Un projet ponctuel sur la synthénie café/tomate/solanées/arabette mené avec Nestlé et l'université de Cornell est à ce niveau intéressant. Un clone BAC de notre banque caféier, situé sur la même zone du génome que des clones BAC de la tomate et de l'arabette (appelé BAC 19), a été entièrement séquencé et cartographié sur la carte génétique. Des marqueurs microsatellites ont été définis à l'intérieur de ce clone BAC pour réaliser de la cartographie fine. Des analyses de déséquilibre de liaison sur quelques dizaines de Kb seront ainsi possibles. Une vingtaine de marqueurs microsatellites est en cours de cartographie dans ce BAC (voir Figure 7). Enfin, un certain nombre de SSR identifiés dans le promoteur et le gène *susy1* est cartographié. Nous avons donc la possibilité de mener une analyse assez approfondie des déséquilibres de liaison dans des zones d'intérêt du génome de *C. canephora*.

### Principales publications de l'équipe

Bustamante-Porras J, Campa C, Poncet V, Noirot M, **Leroy T**, Hamon S, de Kochko A (2007). Molecular Characterization of an Ethylene Receptor gene (*CcETR1*) in coffee trees, its relationships with fruit development and caffeine content. *Molecular Genetics and Genomics*, 277: 701-712.

Cubry P, Musoli P, Legnaté H, Pot D, De Bellis F, Poncet V, Anthony F, Dufour M, **Leroy T** (2008). Diversity in coffee using SSR markers: structure of the *Coffea* genus and perspectives for breeding. *Genome*, 51: 50-63.

Dufour M, Hamon P, Noirot M, Risterucci AM, Brottier P, Vico V, **Leroy T** (2001). Potential use of SSR markers for *Coffea* spp. genetic mapping. In 19<sup>th</sup> International Colloquium on Coffee Science (ASIC), Trieste, Italy, 2001-05-14/18. CD-ROM. ASIC, Paris, France.

Geromel C, Ferreira LP, Bonatelli, Bottcher A, Pot D, Pereira LFP, **Leroy T**, Vieira LGE, Mazzafera P, Marraccini P. (2008). Sucrose metabolism in *Coffea racemosa* endosperm during fruit development. *Annals of Applied Biology*, 152: 179-187.

Geromel C, Ferreira LP, Davrieux F, Guyot B, Ribeyre F, dos Santos Scholz MB, Pereira LFP, Vaast P, Pot D, **Leroy T**, Androcioli Filho A, Vieira LGE, Mazzafera P, Marraccini P. (2008). Effects of shade on the development and sugar metabolism of coffee (*Coffea arabica*) fruits. *Plant physiology and Biochemistry*, 46: 569-579.

Geromel C, Ferreira LP, Guerreiro SMC, Cavalari AA, Pot D, Pereira LFP, **Leroy T**, Vieira LGE, Mazzafera P, Marraccini P (2006). Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. *Journal of Experimental Botany*, 57 (12): 3243-3258.

**Leroy T**, Marraccini P, Dufour M, Montagnon C, Lashermes P, Sabau X, Ferreira LP, Jourdan I, Pot D, Andrade AC, Glaszmann JC, Vieira LGE, Piffanelli P (2005). Construction and characterization of a *Coffea canephora* BAC library to study the organization of sucrose biosynthesis genes. *Theoretical and Applied Genetics*, 111: 1032-1041.

Poncet V, Dufour M, Hamon S, Hamon P, de Kochko A, **Leroy T** (2007). Development of genomic microsatellite markers on *Coffea canephora* and their transferability to other coffee species. *Genome*, 50(12) : 1156-1161.

Pot D, Bouchet S, Marraccini P, De Bellis F, Cubry P, Jourdan I, Vieira LGE, Ferreira LP, Musoli P, **Leroy T** (2006). Nucleotide diversity of genes involved in sucrose metabolism. Towards the identification of candidate genes controlling sucrose variability in *coffee* spp. In 21<sup>th</sup> International Conference on Coffee Science (ASIC), Montpellier, France, September 2006, 11-15, pp.679-686. ASIC, Paris, France.

## Conclusions

L'ensemble des travaux présentés s'inscrit dans une démarche d'amélioration des caféiers et des cafés pour la productivité et la qualité des produits afin d'améliorer les revenus des planteurs. L'ensemble des travaux a été réalisé dans le cadre de collaborations avec les pays producteurs de *C. arabica* et *C. canephora*, afin que les résultats soient directement utilisables par les chercheurs partenaires pour l'amélioration de la caféiculture dans leurs pays.

Ainsi, le programme de sélection récurrente réciproque mis en place en Côte d'Ivoire a permis la mise en place de champs semenciers pour la distribution aux planteurs d'hybrides vigoureux et productifs donnant un café de qualité améliorée par rapport aux sélections précédentes. Nos partenaires indiens ont maintenant les outils pour sélectionner des caféiers transgéniques résistants au foreur des troncs. Nos partenaires ougandais ont à leur disposition du matériel végétal original présentant des caractéristiques intéressantes pour la résistance à la trachéomycose. Enfin, les travaux en cours avec les partenaires brésiliens vont nous permettre de définir des zones du génome impliquées dans la définition de la qualité des cafés et ainsi maintenir celle-ci dans les programmes de sélection en cours.

Depuis 2005, un chercheur de l'équipe au Brésil (P. Marraccini) travaille sur les mécanismes de tolérance à la sécheresse des caféiers. Ces études, particulièrement pertinentes pour un pays comme le Brésil où les prévisions de réchauffement climatique sont assez alarmistes, présentent aussi un grand intérêt pour les pays africains comme la RCI ou l'Ouganda. Les travaux ont débuté dans le cadre d'un projet ABC (Agence Brésilienne de Coopération) entre le Cirad et l'EMBRAPA à Brasilia, et aussi dans le cadre d'un projet ATP Cirad (2007-2009) sur la plasticité des plantes pérennes en conditions de sécheresse (pin, eucalyptus et caféier).

Les travaux débutent, et nous avons commencé la cartographie de quelques gènes qui pourraient présenter un intérêt pour la tolérance à la sécheresse des caféiers. Ces travaux donnent une nouvelle impulsion à notre équipe pour nouer également des coopérations nationales et internationales sur un thème très mobilisateur. Un projet ANR est en cours de rédaction sur le sujet avec des collègues du Brésil, de l'INRA, de l'IRD et de l'Université. Le projet de recherches présenté dans ce document est connecté avec ces travaux en cours et les projets en cours de réflexion.

Un certain nombre d'articles ou d'ouvrages de synthèse ont été publiés sur nos travaux pendant ces 20 ans, sur l'amélioration des caféiers (2002 et 2004), et sur la génétique de la qualité des cafés (2006) :

**Leroy T**, Ribeyre F, Bertrand B, Charmetant P, Dufour M, Montagnon C, Marraccini P, Pot D (2006). Genetics of coffee quality. *Brazilian Journal of Plant Physiology*, 18(1): 229-242.

Eskes AB, **Leroy T** (2004). Coffee selection and breeding. In *Coffee: growing, processing, sustainable production*, JN Wintgens (ed.), pp 57-86. Wiley-VCH Weinheim, Germany.

Montagnon C, **Leroy T**, Bertrand B, Charmetant P, Dufour M (2002). Résultats récents pour l'amélioration génétique du caféier. *Recherche et caféiculture, Plantations, recherche, développement*, ISSN: 1254-7670, pp. 85-94.

## V LE PROJET DE RECHERCHES

### *Génétique d'association chez le caféier *C. canephora**

#### **A. Les questions de recherche**

Chez les espèces végétales, et plus particulièrement pour les plantes pérennes, les résultats de cartographie de QTL ou de gènes candidats montrent que le positionnement de zones précises impliquées dans la variation des caractères d'intérêt est difficile. Comme le montrent nos premiers résultats et ceux d'autres espèces pérennes, l'intervalle de confiance des zones QTL peut dépasser plusieurs dizaines de cM. La réduction de ces intervalles de confiance entre marqueurs et caractère est indispensable pour mettre en place une sélection assistée par marqueurs efficace (Butcher and Southerton 2007). D'autre part, les premiers résultats sur plantes pérennes ont été obtenus sur des arbres producteurs de bois comme le pin (Gonzalez-Martinez et al. 2007), peu de données existent pour des plantes pérennes tropicales à cycle long produisant des fruits.

Notre projet permettra de contribuer à répondre à la question scientifique globale du déterminisme génétique de la tolérance à la sécheresse et de la construction de la qualité en conditions de stress pour des populations sauvages et cultivées chez une plante pérenne tropicale à fruits, le caféier *C. canephora*.

Pour mettre en place ce projet sur le caféier *C. canephora*, les questions scientifiques suivantes devront être traitées:

- Quels sont la structure des populations et les patrons de déséquilibre de liaison dans les différentes entités mises en évidence dans les études de diversité?
- Quelles sont les déterminismes génétiques de la tolérance au stress hydrique et de la qualité : variabilité phénotypique, effet de l'environnement, de l'âge et interactions ?
- Quelles sont les zones chromosomiques d'intérêt pour ces caractères ?
- Quels sont les gènes candidats impliqués (études d'expression sur des génotypes contrastés) ?
- Quelles sont les relations entre ces gènes candidats identifiés et la variabilité phénotypique des caractères d'intérêt au sein des populations ?
- Comment mettre en place une SAM (Sélection Assistée par Marqueurs) efficace sur caféier ?

#### **B. Contexte - L'amélioration des caféiers**

L'amélioration du caféier *C. canephora*, espèce diploïde et allogame, s'est faite dans différents centres de recherche au cours du vingtième siècle, en particulier en Afrique centrale (centres de recherches de Yangambi, en RDC et de Boukoko en RCA) puis dans différents centres d'Afrique occidentale (Côte d'Ivoire, Cameroun) ou d'Afrique orientale (Ouganda). Depuis les années 50, des programmes se sont développés essentiellement pour l'amélioration de la productivité des caféiers. L'amélioration a tout d'abord été basée sur la sélection de familles de plein frères ou de demi-frères. Des essais de familles de demi-frères ont permis la

mise en place de champs semenciers tri clonaux, en particulier en Côte d'Ivoire et au Cameroun. Cependant, le caractère strictement allogame du caféier *C. canephora* et son haut niveau d'hétérozygotie induisent une très grande variabilité génétique dans les descendance. Une sélection de clones dans les meilleures descendance a donc été effectuée, et des clones sélectionnés ont été proposés aux planteurs de la plupart des pays africains producteurs (Côte d'Ivoire, Togo, Zaïre, RCA, Cameroun, Ouganda). Depuis le début du 20<sup>ème</sup> siècle, des échanges de matériel végétal très intenses ont eu lieu entre les centres de recherche, et ont permis de mettre en place des collections de travail avec des génotypes locaux et des génotypes introduits d'autres centres de sélection. La station de recherches du CNRA de Côte d'Ivoire, à Divo, possède donc une collection de travail de caféiers cultivés de plus de 600 génotypes locaux ou introduits.

D'autre part, depuis le début des années 70, des prospections de caféiers sauvages ont été effectuées dans une grande partie de l'aire de répartition du caféier *canephora* (Côte d'Ivoire, Guinée, Cameroun, RCA, Congo) sous la Direction d'équipes de l'IRD et du Cirad. Ainsi, en Côte d'Ivoire, sur la station de Divo, des collections de caféiers sauvages de l'espèce regroupent actuellement plus de 800 génotypes sur différentes parcelles de la station, plantées au fur et à mesure du déroulement des prospections.

Nous disposons donc sur la station de deux collections très complètes, l'une de génotypes cultivés, l'autre de caféiers sauvages. Dans les années 80, l'ensemble des collections de caféiers cultivés et sauvages a été utilisé dans le cadre du programme de sélection récurrente débuté en Côte d'Ivoire (cf IV A-1). Les collections de travail sont plantées sur une parcelle unique, mais les caféiers plantés sont d'âge assez divers, puisque les introductions dans ces collections se sont faits en continu depuis la création de la station, en 1962. Chaque génotype est représenté par 4 arbres plantés en ligne, avec deux arbres en croissance libre et deux arbres écimés. Il n'y a donc pas de dispositif statistique pour l'évaluation de ces arbres. Les collections de caféiers sauvages prospectés dans les forêts d'Afrique occidentale et centrale ont été plantées au fur et à mesure du déroulement des prospections, elles sont donc réparties sur plusieurs parcelles d'âge différent. Chaque génotype est représenté par deux arbres en ligne.

Les évaluations phénotypiques de ces collections ont porté essentiellement sur des caractères technologiques et biochimiques : taille des fèves, taux de fèves caracoli, taux de caféine. Nous ne disposons pas de données sur tous les génotypes et il n'existe pas de données prises la même année sur les arbres de toutes les collections. La collection de travail a fait l'objet d'évaluations plus précises, puisque des données sur l'architecture des arbres, des caractéristiques botaniques et des données technologiques et biochimiques sur plusieurs années existent. Cependant, ces données ne sont pas utilisables telles quelles dans des analyses d'association, à cause de l'âge très différents des arbres, et de l'étalement dans le temps de ces notations.

D'autres collections de *C. canephora* existent aussi dans des pays comme l'Ouganda et le Brésil, mais elles ne couvrent pas une variabilité aussi importante de l'espèce. Cependant, elles pourraient être considérées pour des études d'association ciblées.

Avec les évolutions récentes du climat en Afrique et dans toutes les zones intertropicales (GIEC, 2007), la tolérance à la sécheresse doit être prise en compte dans les programmes d'amélioration. Notre projet consiste donc à comprendre les déterminismes génétiques et à identifier les gènes contrôlant des caractères d'intérêt chez *C. canephora*, plante pérenne

tropicale, dans l'objectif de développer des outils pour mettre en place la sélection assistée par marqueurs. Pour répondre aux questions de recherche, nous aborderons successivement les analyses de QTL, les analyses de déséquilibre de liaison, les études d'association proprement dites pour terminer sur la mise en place de la SAM. Les travaux sur les déterminismes génétiques de la qualité et de la tolérance à la sécheresse sont effectués par l'équipe en partenariat avec les collègues brésiliens, ivoiriens et ougandais dans les projets en cours.

### ***C. Analyses de QTL***

Des analyses QTL sont en cours sur la descendance que nous travaillons en Côte d'Ivoire pour des caractères biochimiques de qualité de la boisson (voir IV D-3). Les travaux se poursuivent sur cette descendance, mais les problèmes principaux viennent de la faible résolution de ces analyses et de l'instabilité des QTL au cours du temps. En effet, les analyses sont effectuées sur les plantes pérennes plusieurs années de suite, afin de considérer les effets environnementaux et les conditions de développement qui sont différentes d'une année à l'autre. La recherche de QTL stables dans le temps pour les caractères d'intérêt est un point primordial pour les espèces pérennes (Pot, 2004 ; Bundock et al. 2008).

Les résultats disponibles montrent bien les limites de nos analyses QTL. En particulier, les dispositifs au champ doivent être adaptés pour permettre une évaluation pluriannuelle fiable. En outre, nos analyses de QTL présentent les inconvénients habituellement imputables à de telles études: faible diversité utilisée, malgré l'utilisation d'un pseudo backcross qui maximise cette variabilité, difficultés à généraliser à toute l'espèce les résultats obtenus, difficultés à utiliser ces résultats dans les programmes d'amélioration étant donné la descendance spécifique et les faibles degrés d'explication pour les caractères les plus intéressants. Sur la base de notre population en cours d'analyse, qui est un pseudo back cross sur Guinéen d'un hybride inter groupe Guinéen par Congolais, il sera possible de détecter les QTL expliquant les différences entre les deux groupes. Mais, pour analyser la variabilité intragroupe et faire des analyses des sous-groupes congolais, nous ne disposerons pas de toutes les informations nécessaires. Il faudrait alors développer de nombreuses populations différentes, ou travailler sur des populations d'association comme notre projet le propose.

Les analyses de QTL vont donc se poursuivre sur la descendance de RCI utilisée pour le projet INCO. D'autre part, il serait utile de travailler sur une descendance mieux adaptée à l'analyse des mécanismes de tolérance à la sécheresse. On peut penser à des croisements de type Guinéens (tolérant) X congolais (sensible) recroisé par un génotype tolérant (SG1 congolais). Le nombre de plantes à mettre au champ sera important, et dépendra de l'héritabilité des caractères, de la variabilité observée, et de la précision des résultats souhaitée (Muranty, 1997), mais des descendance de plus de 300 individus sont souhaitables pour des descendance de pleins frères. Ces analyses en cours ou à venir constituent une base pour déterminer des zones du génome potentiellement impliquées dans la variation de caractères d'intérêt, qui pourront être précisées dans les études d'association.

### ***D. Analyses de DL***

Avant de mener des études d'association, la connaissance du déséquilibre de liaison (DL) dans nos populations est indispensable (Flint-Garcia et al. 2003). Les profils de DL dépendent des caractéristiques de l'espèce, et de la structuration de la diversité dans



l'espèce. La connaissance du DL n'est pas forcément indispensable par elle-même, mais dans nos populations extrêmement diversifiées d'un point de vue génétique et dont l'histoire est très différente suivant les populations, c'est une première étape indispensable pour des études d'association efficaces. L'analyse du DL permet de fixer la distance de pertinence des liaisons entre marqueurs et phénotype (Gupta et al. 2005) pour chaque population d'étude. Les analyses du DL ont débuté dans le cadre du DEA de Komlan Avia et de la thèse de Philippe Cubry en cours dans l'équipe. A partir des analyses de diversité récentes sur les caféiers de l'espèce *C. canephora* (Cubry et al. 2007 ; Musoli et al. soumis), un certain nombre de groupes de diversité ont été définies (cf A-1) :

- les 4 groupes déjà identifiés chez les congolais : SG2, SG1, B et C,
- le groupe des congolais ougandais sauvages,
- les deux populations du groupe Guinéen : la population Pélési et l'ensemble des autres origines sauvages et cultivées.

Les analyses de DL peuvent se faire :

- sur des marqueurs microsatellites (SSR) faciles à utiliser, mais qui présentent l'inconvénient de ne pas être bialléliques. Il est donc difficile de déterminer des valeurs de DL précises, et seul un DL global génotypique est déterminé par des études microsatellites. En effet, la difficulté de mise en évidence des haplotypes rend les analyses de DL avec des SSR moins puissantes, en particulier pour mesurer le DL des allèles des marqueurs sur le même brin chromosomique (vrai DL physique);
- Sur des SNP qui seront préférés car ils sont bialléliques et la détermination des haplotypes en est facilitée, permettant ainsi la mise en évidence d'un DL physique. Les SNP sont d'autant plus pertinents que l'étendue du DL est faible (Gonzalez-Martinez et al. 2007).

Pour le caféier, les analyses de DL ont débuté avec des marqueurs microsatellites sur deux populations congolaises et deux guinéennes :

- Deux populations naturelles, les caféiers de la Nana (groupe C), congolais à effectif important (91 individus) et structuration faible et les Pélési, guinéens à effectif plus faible (37 individus) mais possédant des caractéristiques phénotypiques particulières,
- Une population composite, les cultivés d'origine congolaise (groupe SG2, 83 individus),
- Un mélange de 5 populations naturelles et 1 cultivée (Guinéens autres que Pélési, 129 individus).

L'utilisation des SNP va débiter pour les analyses de DL. Le travail important de détection de polymorphismes dans des gènes candidats possibles (sucres et lipides) a déjà été effectué dans notre équipe, et nous disposons donc de polymorphismes dans un échantillon de 36 génotypes de l'espèce, soit un échantillon représentatif de la diversité de l'espèce (cf IV D-4). Pour les gènes candidats de tolérance à la sécheresse en cours d'analyse, la même recherche de polymorphismes sera effectuée.

L'intensité du DL peut être variable sur les différents chromosomes d'une plante ou sur différentes parties d'un même chromosome (Brescaglio et Sorrells, 2006). De même, on observe souvent des valeurs de DL significatives entre les chromosomes, mais il est possible de prendre en compte ce background pour les analyses, afin d'estimer aussi précisément que

possible le DL entre marqueurs sur le même chromosome. Ces analyses de DL sont donc un préalable indispensable à des études d'association pour déterminer la densité de marquage nécessaire, et l'efficacité possible de ces études. En effet, un DL étendu nécessite un marquage moins dense, mais donne une résolution plus faible dans les analyses d'association. Une décroissance rapide du DL est une situation favorable pour les études d'association au niveau du gène candidat (Szalma et al. 2005), mais ceci implique un marquage plus dense.

Enfin, les études d'expression en cours, pour le projet Génoplante sur la qualité des cafés ou pour celles qui sont prévues dans le cadre du projet ATP « plasticité des plantes pérennes » devraient également nous fournir des informations intéressants. Ces projets font intervenir des génotypes contrastés, et devraient nous permettre d'identifier des gènes candidats pour la qualité et la tolérance à la sécheresse à travailler dans nos études d'association.

## ***E. Etudes d'association***

Les études d'association ont débuté dans le domaine de la santé humaine, avant d'être entreprises sur les plantes, puis beaucoup plus récemment sur les plantes pérennes. Des résultats intéressants ont été obtenus sur le blé avec le choix précis de gènes candidats pour un caractère donné (Ravel et al. 2006) en classant les SNP des différents gènes en fonction du phénotype. Une revue des résultats sur les arbres a été publiée récemment (Neale, 2007). Sur les plantes pérennes, l'impossibilité de réaliser des lignées recombinantes, ou la difficulté de créer des populations multiples, renforce l'intérêt des études d'association, qui analysent des populations larges de plantes non apparentées.

Ces études peuvent se faire à plusieurs niveaux, du polymorphisme mis en évidence sur un gène particulier à une étude sur l'ensemble du génome (whole genome scanning) en passant par des travaux sur un gène candidat donné ou de la cartographie fine de zones particulières du génome (Balding, 2006 ; Burke et al. 2007). Ces études sont complémentaires avec les analyses QTL classiques (Szalma et al. 2005) et évitent de mettre en évidence les associations spécifiques entre marqueurs et QTL dans des populations biparentales (Wei et al. 2006).

Les études d'association présentent en particulier un grand intérêt pour :

- Travailler sur une plus grande variabilité génotypique et phénotypique que dans les études de cartographie classiques, et travailler spécifiquement sur les populations d'amélioration (Gonzalez-Martinez et al. 2007),
- Travailler sur un grand nombre de gènes simultanément (Butcher et Southerton, 2007),
- Travailler plus facilement sur des caractères complexes (Gonzalez-Martinez et al. 2007),
- Travailler en maintenant une base génétique large dans les études (Butcher et Southerton, 2007).

### **1 Quelles populations de travail?**

Sur le caféier, comme sur toutes les plantes, il est fondamental de connaître la structure des populations à étudier (Pritchard, 2001, Butcher et Southerton, 2007). En effet, l'éventuelle structure des populations induit des biais dans les résultats des études d'association, et il est donc préférable de travailler sur des populations non structurées d'un point de vue génétique,

ou dont la structure est connue et peut être prise en compte dans les analyses (Pritchard et al. 2000).

D'autre part, le DL, comme nous l'avons déjà observé sur le caféier, peut également être différent suivant les populations, par exemple à cause de phénomènes de « bottleneck », comme nous l'avons mis en évidence pour la population guinéenne Pélési, constituée vraisemblablement à partir de deux pieds mère. Si l'on travaille sur des populations de petite taille ou soumises à des sélections, celles-ci peuvent aussi présenter des valeurs de DL fortes sur des zones particulières (Rafalski et Morgante, 2004).

Les études menées sur les caféiers depuis quelques années avec des marqueurs isozymes, RFLP puis avec des marqueurs microsatellites ont permis de définir des populations d'études pour l'analyse du DL. Ces populations peuvent être utilisées en tant que telles pour les études d'association. Cependant, il est aussi possible de créer des populations « synthétiques », résultats d'un mélange de plantes issues de plusieurs populations sous forme d'une « core-collection » non structurée, adaptée pour les études d'association (Neale et Savolainen, 2004 ; Gupta et al. 2005). Ainsi, il faut connecter nos études d'association au programme de sélection récurrente en cours en Côte d'Ivoire depuis 25 ans. Ce programme reposant sur la complémentarité phénotypique des groupes Guinéens et Congolais, on peut aussi envisager de créer des populations d'association guinéennes et congolaises, regroupant des génotypes des différentes sous populations mises en évidence, sur une base phénotypique et moléculaire, comme le présentent Breseghello et Sorrells (2006) .

La taille des échantillons pour les analyses est importante pour augmenter le pouvoir de détection et quantifier l'effet d'un plus grand nombre d'allèles (Breseghello et Sorrells, 2006). Pour les caféiers, même si l'hétérozygotie est très forte dans tous les génotypes, il sera préférable de travailler sur des populations de grande taille. Des analyses sur des populations plus restreintes comme les guinéens de Pélési (50 individus) ou les caféiers SG1 (une trentaine d'individus) sont toutefois possibles si l'on conserve un nombre d'allèles important. Dans tous les cas, il faudra s'assurer par les études de DL préliminaires de l'homogénéité du déséquilibre dans les populations créées pour les études d'association. On peut donc travailler sur différentes populations homogènes avec des données de DL connues, mais on peut aussi travailler sur des « populations composites » en tenant compte de la structure.

Dans les études d'association, il peut aussi exister un problème de fausses associations mises en évidence à cause de cette structuration, à des bottleneck sur certaines populations (cas de la population Pélési), à une sélection naturelle dans certaines populations, à des hybridations entre des populations (cas des génotypes N'Ganda et Erect en Ouganda) ou à de la dérive génétique, toutes choses qui ont pu avoir lieu dans nos populations (Ochieng et al. 2007). Pour remédier à ces problèmes de structuration, il faudra allouer les individus au niveau le plus fin possible de population, en traitant à part les hybrides entre les populations, en considérant comme covariable la proportion évaluée de chaque ancêtre.

Pour les caféiers, les résultats en cours d'acquisition sur les DL dans différentes populations nous permettront donc de définir des populations naturelles ou synthétiques adaptées à des études d'association, qui devront comporter un nombre d'individus suffisant. Au moins 200 génotypes seraient nécessaires si l'on se base sur des études déjà réalisées, sur le pin plus de 400 génotypes ont été travaillés dans ces études (Gonzalez-Martinez et al. 2007). Les populations choisies doivent correspondre à des unités de sélection pour le programme d'amélioration en cours. Nous pouvons donc proposer de travailler sur :

- Une population de congolais de différentes origines (SG1, SG2, B et C),
- Une population de génotypes Guinéens issus des différentes populations sauvages et cultivées prospectés en plantation,
- Une population de génotypes ougandais.

Ces populations correspondent aux grandes populations d'amélioration travaillées et permettront de disposer d'un nombre important de génotypes pour chacune d'elles. Dans un deuxième temps, il pourrait être intéressant de travailler sur les différents sous-groupes congolais qui sont considérés maintenant séparément dans le programme d'amélioration et présentent des caractéristiques différentes. De même, les Guinéens cultivés, qui ont en général une meilleure valeur phénotypique propre que les génotypes sauvages, et couvrent l'ensemble de la diversité du groupe, seraient intéressants à considérer séparément.

L'utilisation d'une « core collection » globale de *C. canephora* peut être intéressante d'un point de vue scientifique pour une approche globale des associations dans l'espèce. De plus, pour l'analyse de certains caractères, la disponibilité dans les mêmes parcelles de génotypes contrastés sera très utile.

## 2 Evaluation génotypique des collections

Pour le caféier, se pose le problème de l'utilisation des marqueurs microsatellites ou des SNP pour effectuer le marquage des génotypes. Comme nous l'avons déjà vu, les microsatellites présentent l'inconvénient de leur multi allélisme, mais ils sont peu coûteux et faciles à travailler. La densification de la carte génétique va se poursuivre avec des marqueurs SSR neutres ou liés à des gènes d'intérêt (marqueurs dans les gènes ou dans des séquences terminales de clones BAC). Il faudra aussi travailler avec des marqueurs SNP. Le coût du génotypage SNP a beaucoup diminué ces dernières années, ce qui rend cette technologie plus facile à mettre en œuvre, que ce soit en utilisant les outils de la PCR quantitative ou le séquençage direct.

Il existe des stratégies de tri de la diversité qui permettraient de disposer rapidement d'un grand nombre de polymorphismes (Whetten and Frampton, 2008). On peut concevoir de travailler à partir d'une vingtaine de génotypes, en prenant des ARN à des stades clés pour la réaction des plantes en conditions de sécheresse ou les fruits en cours de développement. On peut ensuite faire un séquençage systématique, qui permet de déterminer à la fois les séquences les plus exprimées, et la diversité nucléotidique de ces gènes pour les génotypes utilisés.

Notre étude pourrait se faire sur les zones suivantes :

- Couverture globale du génome,
- zones QTL repérées pour l'acidité de la boisson, la taille des fèves, la productivité cumulée, la tolérance à la sécheresse ou des caractères biochimiques de qualité de la boisson (caféine, saccharose, lipides, acides chlorogéniques),
- éventuellement, des zones où sont localisés des gènes d'intérêt, s'ils sont colocalisés avec des QTL.

La densité de marquage sera définie en fonction de l'étendue du DL observé, et pourra être différente suivant les populations.

### **3 Analyses phénotypiques**

L'évaluation phénotypique des plantes est le point clé des études d'association (Neale et Savolainen, 2004). Sur les plantes pérennes comme le caféier, c'est le point le plus difficile pour construire des études d'association d'intérêt pour l'amélioration.

En effet, pour que les données soient utilisables, il faut que toutes les plantes soient dans un environnement comparable ou contrôlé, au même stade physiologique. L'évaluation phénotypique en conditions totalement contrôlées est possible pour des plantes annuelles ou des travaux sur un stade particulier (croissance foliaire ou racinaire par exemple pour une tolérance à la sécheresse) pour les plantes pérennes, ce qui exclut de pouvoir travailler sur des caractères complexes comme la productivité ou la qualité des fruits. L'installation d'essais en champ est alors indispensable, même s'il est extrêmement difficile et coûteux de mettre en place des plans d'expérience aussi complets et performants (nombre de répétitions, essais multi locaux) que sur les plantes annuelles.

En théorie, comme pour certaines plantes, des données phénotypiques existent déjà et sont utilisables, ce qui réduirait les coûts (Wei et al. 2006). Pour les caféiers, la multiplication des parcelles des collections, l'âge très variable des plantes et l'extrême hétérogénéité des sols latéritiques rend cette utilisation difficile. Il faudra donc mettre en place des parcelles spécifiques destinées à évaluer les phénotypes des génotypes analysés. Pour prendre en compte l'hétérogénéité des terres, nous proposons un dispositif en randomisation totale arbre par arbre, avec au minimum 5 arbres par génotype étudié. Ce dispositif, lourd pour l'analyse de plusieurs centaines d'arbres, est le seul possible pour analyser correctement le phénotype des plantes. Un « effet champ » existera sur les données phénotypiques, il sera très difficile à prendre en compte et pourrait avoir des conséquences importantes (Burdon et Wilcox, 2007). C'est pourquoi le dispositif doit être aussi complet que possible, dans des conditions de champ aussi homogènes que possible.

Les caractères analysés porteront essentiellement sur la productivité des arbres et les facteurs de production, la qualité des fruits (taille des fèves, taux de caracoli) et les caractéristiques biochimiques et organoleptiques du café produit. Enfin, l'accent sera mis sur la tolérance à la sécheresse des arbres jeunes (installation et mortalité) et adultes (capacité à produire en conditions de sécheresse).

### **4 Analyse des associations**

L'analyse des associations entre phénotype et génotype peut se faire par des analyses de variances (modèle linéaire généralisé, MLG) ou par des études de régression (Wei et al. 2006), en particulier pour les analyses de maladies ou des notations qualitatives.

Les faux positifs possibles, dus à la structuration des populations, contrôlée ou non, pourront être limités en utilisant la structure des populations comme covariable dans nos analyses (Brescaghiello et Sorrells, 2006) ou en utilisant un MLG à effets mixtes, avec le marqueur en effets fixes et la structuration de la population en effet aléatoire.

L'existence d'interactions épistasiques pourrait aussi compliquer nos analyses. En effet, l'expression phénotypique d'un allèle peut être modifiée par des allèles d'un autre locus. Cette influence du background génétique, caractérisée par la présence d'un DL non physique entre les marqueurs sur le génome souligne l'importance de travailler en multi locus pour un caractère considéré (Szalma et al. 2005). Il existe des possibilités de modélisation de ce DL non physique (Jannink, 2006) si cela s'avère nécessaire.

Ce programme d'étude d'associations sera développé avec les collègues des pays producteurs, en particulier avec les ivoiriens qui accueillent les collections les plus importantes. Le coût important du génotypage, même s'il est en baisse, sera à prendre en compte dans le déroulement de nos études. Mais, le coût principal de nos études sera dû à l'analyse phénotypique qui nécessitera la mise en place de champs de plusieurs centaines d'individus à observer individuellement pendant plusieurs années, avec des évaluations de caractères multiples allant du développement végétatif à la caféine dans les grains. Selon Butcher et Southerton (2007) sur les arbres forestiers, il faudrait des essais de 500 individus du même âge sur le même site pour arriver à mettre en évidence des nucléotides responsables de 5% de la variation phénotypique sur des arbres. Il n'est pas sûr qu'une résolution aussi fine soit indispensable sur le caféier, mais cela illustre bien les contraintes propres aux plantes pérennes.

## ***F. Perspectives de sujets de recherche***

Le projet de recherches présenté permettra d'aborder un certain nombre de sujets d'intérêt scientifique. Tout d'abord, même s'il est prévu de créer de nouvelles parcelles pour travailler sur les études d'association, il ne sera pas possible de mesurer tous les caractères sur tous les arbres dans les mêmes conditions la même année, que ce soit pour des données sur l'architecture des arbres, leur tolérance aux stress biotiques ou abiotiques, ou les caractéristiques du produit. Il serait donc intéressant de mener un travail de recherches sur la prise en compte dans nos analyses statistiques de l'effet environnemental et de l'effet année, dont on connaît déjà l'importance dans les dispositifs au champ sur les plantes pérennes à fruit en milieu tropical.

D'autre part, en relation avec la tolérance à la sécheresse, deux questions scientifiques retiennent plus spécialement notre attention pour les études d'association. La première question porte sur l'installation des plantules au champ la première année (mortalité, reprise, vigueur), et la définition d'indicateurs phénotypiques et génétiques de cette installation. L'autre sujet d'intérêt concerne le système racinaire des plantes au champ. Les caféiers possèdent un système racinaire superficiel pour les éléments nutritifs et un système profond pour l'eau. Comme ce qui est travaillé sur le riz, il serait intéressant de définir puis de cloner des QTL de développement des racines, éventuellement dans des conditions de stress importantes, qui seraient corrélées à la capacité des plantes à aller chercher l'eau en profondeur. On peut proposer de mettre en place un essai spécifique sur les stress des caféiers au champ sur un cycle de sélection (5 ans). Cela permettrait d'analyser l'installation des plantes, leur tolérance à tous les stades (floraison, différents stades du développement des fruits, stress en présence de stress biotiques) à des stress de sécheresse dans des conditions de champ. Cet essai devrait comporter des génotypes connus comme tolérants (Guinéens) ou sensibles (Congolais SG2), mais aussi des structures génétiques hybrides très productives. Au delà de l'aspect sécheresse, un tel essai permettra d'analyser comparativement les qualités des

cafés produits par les arbres en analysant l'expression des gènes impliqués dans la construction de la qualité.

Pour la qualité des cafés produits, il serait aussi intéressant d'approfondir les études QTL puis éventuellement la mise en évidence d'allèles favorables pour les déterminants des caractéristiques organoleptiques des caféiers : acidité, amertume, astringence et force. Les premiers résultats montrent que cette approche est pertinente. Des travaux sur des populations contrastées (guinéens « mauvais » et congolais « bons » par exemple) avec des protocoles adaptés de prélèvement des échantillons pourraient nous permettre de progresser dans la connaissance de la mise en place de la qualité chez les caféiers.

Enfin, les études d'association sont un élément dans la mise en place d'une sélection assistée par marqueurs pour les caféiers. Les résultats des études d'association seront pris en compte et intégrés dans le programme d'amélioration par sélection récurrente du caféier *C. canephora* en cours en Côte d'Ivoire, avec des brassages intragroupes à la fin de chaque cycle de sélection. Les relations entre phénotype et marqueurs doivent être très fortes, pour ne pas être rompues lors des cycles de recombinaison propres aux cycles de sélection récurrente (Neale, 2007). Il est probable que certains allèles favorables soient fixés dans certaines populations (par exemple pour la tolérance à la sécheresse des guinéens), et ceci devra être pris en compte lors des cycles de sélection. Si les allèles favorables sont fixés dans certaines populations, il sera alors utile de travailler sur des populations de cartographie classique de type rétro croisement ou F2 pour mettre en évidence les allèles favorables dans la descendance. La sélection des caféiers devra donc associer cette sélection par marqueurs avec une sélection phénotypique par index, comme celle qui est déjà en place en Côte d'Ivoire et qui a permis une hausse de productivité de plus de 40% en dix ans de travaux.

A plus long terme, il sera intéressant de voir comment transférer les résultats obtenus sur l'espèce *C. canephora* sur l'autre espèce de caféier cultivé, *C. arabica*. Celle-ci est un allo tétraploïde, dont un des deux génomes d'origine est très proche de l'espèce *C. canephora* actuelle. Cependant, nous ignorons le mécanisme de la spéciation, et nous ne connaissons donc pas le pourcentage de variabilité de *canephora* encore présent chez l'Arabica. Cette espèce présente une variabilité génétique beaucoup plus faible que *C. canephora*, avec une structuration génétique très peu marquée et des collections de caféiers sauvages qui ont été peu travaillées.

Les résultats acquis sur l'espèce diploïde sur les zones du génome les plus pertinentes seront sans doute difficilement transférables sur cette espèce, en raison de la faible variabilité résiduelle. La difficulté viendra aussi du travail de génotypage. En effet, les marqueurs microsatellites sont très peu variables dans cette espèce, et le nombre de SNP mis en évidence est moindre que pour *C. canephora*. D'autre part, le caractère allo tétraploïde de ce caféier rendra l'évaluation du DL difficile. En effet, il pourra exister du DL propre à chaque génome mère et du DL inter génome. Cependant, le lancement de tels travaux serait intéressant, la tolérance à la sécheresse et la qualité du produit seront les facteurs les plus importants à analyser dans les populations. Ces analyses pourraient se faire avec les collègues brésiliens du IAPAR qui possèdent des collections assez importantes qui sont en voie de réhabilitation pour effectuer des analyses phénotypiques indispensables pour les études d'association.

## Références bibliographiques citées dans le document

Anthony F, Lashermes P (2005). Origin, evolution and diversity of the coffee (*Coffea arabica* L.) genome. In Plant genome biodiversity and evolution, Volume 1, Part B Phanerogams, AK Sharma, A Sharma (Eds), pp. 207-228. Science Publishers, Enfield, Great Britain.

Anthony F, Bertrand B, Quiros O, Wilches A, Lashermes P, Berthaud J, Charrier A (2001). Genetic diversity of wild coffee (*Coffea arabica* L) using molecular markers. *Euphytica*, 118 : 53-65.

Balding DJ (2006). A tutorial on statistical methods for population association studies. *Nature review in genetics*, 7:781-791.

Baud S, Vaultier MN, Rochat C (2004). Structure and expression profile of the sucrose synthase multigene family in *Arabidopsis*. *Journal of Experimental Botany*, 55: 397-409.

Berthaud J (1986). Les ressources génétiques pour l'amélioration des caféiers africains diploïdes. Evaluation de la richesse génétique des populations sylvestres et de ses mécanismes organisateurs. Conséquences pour l'application. Document ORSTOM, 379 p. ORSTOM, Paris, France.

Bertrand B, Guyot B, Anthony F, Lashermes P (2003). Impact of *Coffea canephora* gene introgression on beverage quality of *C. arabica*. *Theoretical and Applied Genetics*, 107: 387-394.

Bouharmont P, Lotode R, Awemo J et Castaing X (1986). La sélection générative du caféier Robusta au Cameroun. Analyse des résultats d'un essai d'hybrides diallèle partiel implanté en 1973. *Café, Cacao, Thé*, 30 : 93-112.

Breseghello F and Sorrells ME (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics*, 172: 1165-1177.

Bundock PC, Potts BM and Vaillancourt RE (2008). Detection and stability of quantitative trait loci (QTL) in *Eucalyptus globulus*. *Tree Genetics & Genomes*, 4:85-95.

Burdon RD and Wilcox PL (2007). Population management: potential impact of advances in genomics. *New forests*, 34: 187-206.

Burke JM, Burger JC and Chapman MA (2007). Crop evolution: from genetics to genomics. *Current opinion in genetics*, 17: 525-532.

Butcher P and Southerton S (2007). Marker-assisted selection in forestry species. In *Marker-assisted selection, current status and future perspectives in crops, livestock, forestry and fish*, EP Guimaraes, J Ruane, BD Scherf, A Sonnino and JD Dargie (Eds), pp 283-305. FAO, Rome, Italy.

Capot J (1977). L'amélioration du caféier Robusta en Côte d'Ivoire. *Café Cacao Thé*, 21 : 233-244.

Charrier A (1978). Etude de la structure et de la variabilité génétique des caféiers. *Bulletin IFCC n°14*, 100p. IFCC, Paris, France.

Charrier A et Berthaud J (1975). Variation de la teneur en caféine dans le genre *Coffea*. *Café, Cacao, Thé*, 19: 251-264.

Chevalier A (1947). Les caféiers du globe. III. Systématique des caféiers et faux caféiers. *Maladies et insectes nuisibles*, 365p. Paul Lechevalier, Paris, France.

Combes MC, Andrzejewski S, Anthony F, Bertrand B, Rovelli P, Graziosi G, and Lashermes P (2000). Characterization of microsatellite loci in *Coffea arabica* and related coffee species. *Molecular Ecology*, 9(8): 1178-1180.



- Cordier L (1961). Les objectifs de la sélection caféière en Côte d'Ivoire. *Café, Cacao, Thé*, 5: 147-158.
- Cramer PJS (1957). A review of literature of coffee research in Indonesia, 262 p. SIC Editorial Interamerican Institute of Agricultural Sciences, Turrialba, Costa Rica.
- De Namur C, Couturon E, Sita P et Anthony F (1988). Résultats d'une mission de prospection de cafés sauvages au Congo. In 12<sup>th</sup> International colloquium on Coffee Science, Montreux, Switzerland, pp. 397-404. ASIC, Paris, France.
- Dussert D, Lashermes P, Anthony F, Montagnon C, Trouslot P, Combes MC, Berthaud J, Noirot M et Hamon S (1999). Le caféier, *Coffea canephora*. In Diversité génétique des plantes tropicales cultivées, P Hamon, M Seguin, X Perrier et JC Glaszmann (Eds), pp. 175-194. CIRAD, Montpellier, France.
- Flint-Garcia SA, Thornsberry JM and Buckler ES (2003). Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology*, 54: 357-374.
- GIEC, 2007. Changements climatiques : les bases scientifiques physiques.
- Gonzalez-Martinez SC, Wheeler NC, Ersoz E, Nelson CD and Neale DB (2007) Association genetics in *Pinus tadea* L. I. Wood property traits. *Genetics*, 175: 399-409.
- Guerreiro O, Denolf P, Peferoen M, Decazy B, Eskes AB, Frutos R (1998). Susceptibility of the coffee leaf miner (*Perileucoptera* spp.) to *Bacillus thuringiensis*  $\delta$ -endotoxins: a model for transgenic perennial crops resistant to endocarpic insects. *Current Microbiology*, 36:175-179.
- Gupta PK, Rustgi S and Kulwal PL (2005). Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Molecular Biology*, 57: 461-485.
- Jannink J-L (2007). Identifying quantitative trait locus by genetic background interactions in association studies. *Genetics*, 176: 553-561.
- Jolly D, Taylor D, Marchant R, Hamilton A, Bonnefille R, Buchet G, and Riollet G (1997). Vegetation dynamics in central Africa since 18,000 yr BP: pollen records from the interlacustrine highlands of Burundi, Rwanda and western Uganda. *Journal of Biogeography*, 24: 495-512.
- Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M, Charrier A (2001). Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. *Genome*, 44: 589-596.
- Lashermes P, Combes MC, Trouslot P et Charrier A (1997). Phylogenetic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. *Theoretical and Applied Genetics*, 94: 947-953.
- Maley J (1996). The African rain forest - main characteristics of changes in Vegetation and climate from the upper Cretaceous to the Quaternary. In *Essays on the Ecology of the Guinea-Congo Rain Forest*, IJ Alexander, MD Swine and R Watling (Eds), pp. 31-73. Royal Society of Edinburgh, Edinburgh, Scotland.
- Meunier J et Gascon JP (1972). Le schéma général d'amélioration du palmier à huile à l'IRHO. *Oléagineux*, 27: 1-12.
- Millot F (1974). Recherches en sélection caféière en Ouganda. In *Contribution à l'étude de la cacaoculture et de la caféiculture en Ouganda*, Bulletin IFCC n°12, pp. 47-66. IFCC, Paris, France.
- Montagnon C et Bouharmont P (1996). Multivariate analysis of phenotypic diversity of *Coffea arabica*. *Genetic Resources And Crop Evolution*, 43: 221-227.
- Muranty H (1996). Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity*, 76: 156-165.
- Neale DB (2007). Genomics to tree breeding and forest health. *Current opinion in Genetics*, 17: 539-544.

Neale DB and Savolainen O (2004). Association genetics of complex traits in conifers. *Trends in Plant Science*, 9(7): 325-330.

Noir S, Patheyron S, Combes MC, Lashermes P, Chalhoub B (2004). Construction and characterization of a BAC library for genome analysis of the allotetraploid coffee species (*Coffea arabica* L.). *Theoretical and Applied Genetics*, 109: 225-230.

Ochieng JW, Muigai AWT and Ude GN (2007). Localizing genes using linkage disequilibrium in plants : integrating lessons from the medical genetics. *African Journal of Biotechnology*, 6(6): 650-657.

Portères R (1937). Etude sur les caféiers spontanés de la section *Eucoffea*. Leur répartition, leur habitat, leur mise en culture et leur sélection en Côte d'Ivoire (Répartition et habitat). *Annales de l'Afrique Occidentale Française et Etrangère*, 1 : 68-91.

Portères R (1959). Valeur agronomique des caféiers des types Kouillou et Robusta cultivés en Côte d'Ivoire. *Café, Cacao, Thé*, 3: 3-13.

Pot D (2004). Déterminisme génétique de la qualité du bois chez le pin maritime: du phenotype aux gènes. Thèse ENSA Rennes, 335p.

Pritchard JK (2001) Deconstructing maize population structure. *Nature genetics*, 28: 203-204.

Pritchard JK, Stephens M and Donnelly P (2000). Association mapping in structured populations. *Annual Journal of Human genetics*, 67: 170-181.

Rafalski A and Morgante M (2004). Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends in genetics*, 20 (2): 103-111.

Ravel C, Praud S, Murigneux A, Linossier L, Dardevet M, Balfourier F, Dufour P, Brunel D and Charmet G (2006). Identification of *Glu-B1-1* as a candidate gene for the quantity of high-molecular-weight glutenin in bread wheat (*Triticum aestivum* L.) by means of an association study. *Theoretical and Applied Genetics*, 112: 738-743.

Rovelli P, Mettullo R, Anthony F, Anzueto F and Lashermes P (2000). Microsatellites in *Coffea arabica* L. In *Coffee biotechnology and quality*, T Sera, CR Soccol, A Pandey and S Roussos (Eds), pp. 123-133. Kluwer Academic Publishers, The Netherlands.

Sardana R, Dukiandjiev S, Giband M, Cheng XY, Cowan K, Sauder C, Altosaar I (1996). Construction and rapid testing of synthetic and modified toxin gene sequences CryIA (b&c) by expression in maize endosperm culture. *Plant Cell Reports*, 15: 677-681.

Spiral J, Thierry C, Paillard M, Pétiard V (1993). Obtention de plantules de *Coffea canephora* Pierre (Robusta) transformées par *Agrobacterium rhizogenes*. *Compte-Rendu de l'Académie des Sciences, Paris*, 316 (série III): 1-6.

Szalma SJ, Buckler ES, Snook ME, McMullen MD (2005) Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. *Theoretical and Applied Genetics*, 110: 1324-1333.

Thomas OAS (1940). Robusta Coffee. In *Agriculture in Uganda*, JD Tothil (ed.), pp. 289-315. Oxford University Press, London, Great Britain.

Thomas OAS (1944). The wild coffee of Uganda. *Empire Journal of Experimental Agriculture*, 12: 1-12.

Van Bostel J, Eskes A, Berthouly M (1997). Glufosinate as an efficient inhibitor of callus proliferation in coffee tissue. *In Vitro Cellular and Developmental Biology of Plants*, 33: 6-12.

Wei X, Jackson PA, McIntyre CL, Aitken KS and Croft B (2006). Association between DNA markers and resistance to diseases in sugarcane and effects of population substructure. *Theoretical and Applied Genetics* 114: 155-164.

Whetten RW and Frampton J (2008). Discovering genes and polymorphisms in Fraser Fir by parallel sequencing of PCR-based libraries. In PAG XVI, 2008, January, 14-18, San Diego, poster.

White F (1979). The Guineo-Congolian Region and its relationships to other phytochoria. Bulletin du Jardin Botanique National Belge, 49 : 11-55.